

Mentalizing and Epistemic Trust

Mentalizing and Epistemic Trust

*The work of Peter Fonagy and colleagues at the
Anna Freud Centre*

ROBBIE DUSCHINSKY AND SARAH FOSTER

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Robbie Duschinsky and Sarah Foster 2021

The moral rights of the authors have been asserted

First Edition published in 2021

Impression: 1

Some rights reserved. No part of this publication may be reproduced, stored in a retrieval
system, or transmitted, in any form or by any means, for commercial purposes, without the prior
permission in writing of Oxford University Press, or as expressly permitted by law, by licence or
under terms agreed with the appropriate reprographics rights organization.



This is an open access publication, available online and distributed under the terms of a
Creative Commons Attribution – Non Commercial – No Derivatives 4.0 International licence
(CC BY-NC-ND 4.0), a copy of which is available at <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2020952934

ISBN 978–0–19–887118–7

DOI: 10.1093/med/9780198871187.001.0001

Printed and bound in the UK by
TJ Books Limited

Oxford University Press makes no representation, express or implied, that the drug dosages in this book
are correct. Readers must therefore always check the product information and clinical procedures with
the most up-to-date published product information and data sheets provided by the manufacturers
and the most recent codes of conduct and safety regulations. The authors and the publishers do not
accept responsibility or legal liability for any errors in the text or for the misuse or misapplication of
material in this work. Except where otherwise stated, drug dosages and recommendations are for the
non-pregnant adult who is not breast-feeding

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Acknowledgements

We would like to offer our thanks, above all, to Peter Fonagy, Chloe Campbell, and Liz Allison for their kind and thoughtful feedback throughout our work on this project and on the manuscript. It has meant a huge amount to us. The book has likewise been greatly improved by feedback from Jessica Borelli, Muriel Darmon, Richard Devine, Jeremy Holmes, Elliot Jurist, Alessadro Talia, Everett Waters, Kate White, and David Wilkins.

Special thanks are due to the Wellcome Trust for making it possible for us to work on this book (Grant WT103343MA), and for paying the open access fees.

We are also very grateful to Martin Baum, Janine Fisher, Anne Abel Smith, and colleagues at Oxford University Press for their work on this book.

We feel deeply fortunate for the intellectual companionship and camaraderie of our immediate research group: Lianne Bakkum, Helen Beckwith, Barry Coughlan, Julia Mannes, Stephen Metcalf, Sophie Reijman, Sam Reisz, Guy Skinner, and Melody Turner. We have appreciated all the wonderful conversations we have had, academic and non-academic, with these wonderful people. The support we have received on various projects from Marinus van IJzendoorn, Marian Bakermans-Kranenburg, and Carlo Schuengel has been generous beyond any measure.

Robbie: I am grateful for the incredible support of professional colleagues through the losses and family health challenges of the past few years. My thanks to colleagues in the Primary Care Unit and Sidney Sussex College, Cambridge. It has also been a privilege to work with Julie Anderson and Dustin Hutchinson—colleagues and friends—on the ‘Living Assessments’ project. Finally, I want to thank my wife Sophie, especially for her courage through these difficult years.

Sarah: I would like to thank my PhD supervisors Greta Defeyter and Robbie Duschinsky. I would also like to thank colleagues at Northumbria University, including the social work and guidance and counselling teams, and my office mates. I have been fortunate to benefit from their professional wisdom and support, and also their friendship. My thanks also to my husband Mark for everything.

Finally ... research for this book began six years ago. But writing up coincided with expecting the arrival of our respective daughters, Abigail Duschinsky and Evie Foster. The final revisions and proofs have coincided with the COVID-19 pandemic and lockdowns. Welcome to the world, little Abigail and Evie! When it becomes possible, we look forward to introducing you.

Contents

Introduction	1
Clarification of terms and concepts	3
Central collaborators	7
Overview of the book	9
1. Biographical context	11
Introduction	11
The Fónagy family	11
Fonagy's psychoanalysis with Anne Hurry	15
Mr S	21
2. Work at the Anna Freud Centre	26
Introduction	26
Research director at the Anna Freud Centre	26
The Anna Freud Centre retrospective study	35
Transformations	38
3. Mentalizing in development	49
Introduction	49
Origins of the mentalization concept	50
Reflective function	58
Disorganization and mentalization	69
4. Mentalization in transition	77
Introduction	77
Self-criticism	77
Four poles of mentalizing	82
Defining mentalizing	92
5. Forms of non-mentalizing	102
Introduction	102
Pretend mode	104
Hypermentalizing	106
The capacity for pretend as an asset	111
Psychic equivalence	116
Teleological mode	119
Why three forms of non-mentalizing?	123
6. Conceptualizing the 'self'	133
Introduction	133
What is a self?	136
The primary unconscious	142
Externalizing the alien self	151
Sexuality and aggression	157

7. Adaptation and mental health	166
Introduction	166
Adaptation and attachment	167
Epistemic trust	177
Diagnoses and the p-factor	196
Symptom contexts and networks	209
8. Mentalization-based therapies	216
Introduction	216
The development of mentalization-based therapy	216
Dynamic Interpersonal Therapy	219
Mentalization-Based Therapy today	222
The potential for harm	233
9. The wider social system	238
Introduction	238
The social environment	239
Prevention and public mental health	246
Collective mentalizing capacity	256
Mentalizing cultures	262
Conclusions	282
Constructive appraisal	282
Some remaining questions: mentalizing	285
Some remaining questions: non-mentalizing	287
Some remaining questions: epistemic trust	290
Core strengths of the theory	290
<i>Index</i>	293

Introduction

Peter Fonagy has been described as ‘one of the most acclaimed child psychologists of his generation’, and as leading a ‘revolution’ in the theory and delivery of therapeutic services.¹ He has published over 500 scientific papers and 260 chapters, and has authored or co-authored 19 books. Since 2003, he has served as Chief Executive of the Anna Freud National Centre for Children and Families in London and, since 2008, as Head of the Division of Psychology and Language Sciences at University College London. He has also had a major role in UK health policy, including as Chair of the Outcomes Measurement Reference Group at the Department of Health, Chair of two National Institute for Health and Care Excellence (NICE) Guideline Development Groups, Chair of the Strategy Group for National Occupational Standards for Psychological Therapies, and co-chair of the Department of Health’s Expert Reference Group on Vulnerable Children. In 2015, he was the first UK recipient of the Wiley Prize of the British Academy for Outstanding Achievements in Psychology by an international scholar.

Fonagy’s research has been highly interdisciplinary, moving from psychoanalysis to attachment theory, and from attachment theory to a wider agenda in developmental psychopathology. He has attracted collaborators from among the most talented researchers and clinicians from various disciplines. These collaborations have combined with Fonagy’s remarkable receptivity and desire to learn to contribute to sustained theoretical change and elaboration over the decades. Yet, this continually evolving theory poses difficulties for readers in understanding and applying the approach. Fonagy has joked that readers of a psychological theory can ‘imprint’ on the work of one era, failing to see changes in a theorist’s stance or in the meaning given to terminology.² On the one hand, there may be an undertow from early impressions, which lead later developments to be ignored or downplayed. Equally, someone familiar with the later work of Fonagy and colleagues may miss out on many insights from earlier periods that have not subsequently seen sustained attention. They may also miss the underpinning architecture of the ideas, which comes into focus most clearly when the diverse writings of Fonagy and colleagues are considered together and over time.

The volume and spread of these writings have also, to date, contributed to the lack of an integrative overview and evaluation.³ This will be our central goal here, continuing work in an earlier book—*Cornerstones of Attachment Research*—offering an integrative appraisal

¹ Doward, J. and Hall, S. (2019). ‘Therapy Saved a Refugee Child. Fifty Years on, He’s Leading a Mental Health Revolution’. *Guardian*, 27 April. Accessed at: <https://www.theguardian.com/society/2019/apr/27/peter-fonagy-refugee-child-psychologist-anna-freud-centre>.

² Fonagy, P. (2015). ‘Mutual Regulation, Mentalization, and Therapeutic Action: A Reflection on the Contributions of Ed Tronick to Developmental and Psychotherapeutic Thinking’. *Psychoanalytic Inquiry*, 35(4): 355–369: ‘As with Freud, Bowlby’s critics were often apparently imprinted with the initial model; their attitude of hostility did not permit noting the change in Bowlby’s view’ (p. 357). Perhaps the most widespread case in the reception of the work of Fonagy and colleagues has been that the majority of readers seem to have imprinted on the idea of ‘pre-mentalizing’ states, as a kind of regression to childhood thinking, failing to notice that this has been replaced by the idea of ‘non-mentalizing states’ (see Chapter 5).

³ Liljenfors, R., and Lundh, L. G. (2015). ‘Mentalization and Intersubjectivity towards a Theoretical Integration’. *Psychoanalytic Psychology*, 32(1): 36–60. ‘This theoretical framework, which we will refer to as the theory of mentalization, is probably not yet to be seen as a full-fledged theory. Rather it is as a framework which is still

of the ideas of major psychological theorists who have engaged with ideas of attachment.⁴ *Cornerstones of Attachment Research* presents a comprehensive analysis of the ideas of Bowlby, Ainsworth, Main and Hesse, Sroufe and Egeland, and Shaver and Mikulincer. The present book can certainly be read independently, but—in Chapters 3 and 7 especially—is also a continuation of the ongoing story of attachment research, and its attendant strengths and weaknesses.

The need for an integrative overview and evaluation additionally stems from the segmentation of the audience of the works of Fonagy and colleagues. In a sense, each audience has had direct access to only a ‘part object’, obstructing both their evaluation and their use of the contributions of Fonagy and colleagues. Readers who know the work of Fonagy and colleagues on randomized trials may not know their work on the details of clinical technique, hindering their sense of how the clinical modality actually should work in practice. Readers who know the work of Fonagy and colleagues in attachment theory may have only a general sense of the model of modes of non-mentalizing, not only obscuring the fundamentals of the account of pathology but also of factors that undermine attachment security. Readers who have trained in delivering mentalization-based therapy may nonetheless not know the recent thinking of Fonagy and colleagues about the primary unconscious and, as a result, may push at closed doors with patients, or even deliver therapy in ways that make their patients’ symptoms worse.

We acknowledge that there are limitations to a review of written texts as a methodology for understanding a living psychological theory. Not least, the textual record offers an opaque window on the process by which, and the context within which, scientific and theoretical work was actually completed. Because this process and context shapes the very meaning of the work, this is a significant deficiency. One partial solution is to read the penumbra of wider texts in order to understand the interventions that were intended by a particular work; another partial solution is to supplement the published record with unpublished sources in the public domain, when these are available.⁵ A further strategy is to receive and integrate feedback from the psychological theorists themselves, in this case Fonagy and colleagues. All these solutions have been pursued, but they offer only an incomplete fix. In particular, it should be highlighted that we are not mentalization-based therapists: we have no doubt that this would be a different book, and have different strengths, if we had been trained and were practitioners in this modality.

The reader should also note that the phrase ‘Fonagy and colleagues’, often used in this book, refers to a complex assemblage of people, discourses, institutional structures and priorities, research funding, scientific tools, clinical settings—and their interaction. It is *both* a collective endeavour of thought and effort, and a product of the achievements and trajectory of individual scholars within this group, with different areas of expertise.⁶ The theory is

developing’ (p. 37). In our view, there is no contradiction between the status of full-fledged theory and ongoing development.

⁴ Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press.

⁵ Skinner, Q. (2002). *Visions of Politics: Regarding Method*, Cambridge: Cambridge University Press.

⁶ Griffin, R. J. (2019). ‘The Profession of Authorship’, in Simon Eliot and Jonathan Rose (eds), *A Companion to the History of the Book*, New York: Wiley, pp. 773–785. In a sense, ‘Fonagy’, like the Crown in medieval legal and constitutional theory, has both individual and corporate identities.

‘Fonagy’ refers to an individual person, Peter Fonagy, who holds certain attitudes and has pursued certain actions. Yet, ‘Fonagy’ also refers to the embodiment of the assemblage of theoretical work occurring under his aegis. In this book, the phrase ‘Fonagy and colleagues’ is an attempt to signal the corporate identity, and ‘Peter Fonagy’ is invoked to refer to the individual person. However, there is inevitable instability: this division does not always hold. Maitland, F. W. (2003). *State, Trust and Corporation*, Cambridge: Cambridge University Press.

the result of what Hutchins calls ‘distributed cognition’, a network of processes that together create effects greater than the sum of their parts (see Chapter 7).⁷ This collective contribution has been frequently highlighted by Fonagy: ‘The work summarized in this paper is the result of a collaborative effort of a group of wonderfully talented individuals who have honoured the author with their friendship over the past years.’⁸ He joked in a paper from 2000: ‘This paper is a preliminary report of an ongoing collaboration with friends and colleagues Mary Target, George Gergely, and Efrain Bleiberg. Many of the ideas presented are theirs, but if they should be well received the author will have no hesitation in taking credit for them.’⁹ The differences between the multifaceted labour of clinical and research work and its frontstage, published description is just one of the many limitations of any historical review dependent on texts. However, one opportunity stemming from a comprehensive review of written texts over time is the chance for perspective taking.

Clarification of terms and concepts

Joseph Sandler, Fonagy’s friend and predecessor as Freud Professor at University College London, was very concerned to give theoretical terms due scrutiny and consideration. Sandler emphasized that, just as humans need to engage in reality testing in order to benefit from both fantasy and perception, so psychological theory itself needs to be treated with due seriousness and be subject to ‘concept testing’. Because concepts help guide perception and thought, they will benefit from discussion, appraisal and, at times, ‘adaptive reorganization’ in order to remain in good working condition.¹⁰ He acknowledged that concepts need some elasticity in order to ‘take up the strain of theoretical change, absorbing it while more organized newer theories or part-theories can develop’; yet he also worried ‘concepts become stretched to encompass new insights and new ideas. Often such an expansion of the meaning of a conceptual term is not explicit’, and this can cause confusion and hinder scientific and clinical developments.¹¹ Sandler observed that there can be resistance to efforts to appraise concepts that seem to work well enough for pragmatic purposes, even when it is known that their use nonetheless has resulted in significant disadvantages:

⁷ Hutchins, E. (1995). *Cognition in the Wild*, Cambridge, MA: MIT Press. On psychoanalytic theory as an ‘interactive’ product of individual thinkers and of the collective capacities of communities, see Winnicott, D. W. ([undated] 2016). ‘Outline for a Study in the Sociology of Knowledge’, in Lesley Caldwell and Helen Taylor Robinson (eds), *The Collected Works of D. W. Winnicott*, Volume 9, Oxford: Oxford University Press, pp. 391–394.

⁸ For example, Fonagy, P. and Target, M. (2006). ‘The Mentalization-Focused Approach to Self Pathology’. *Journal of Personality Disorders*, 20: 544–576; George Moran, George Gergely, Miriam and Howard Steele, Helen Stein, John Allen, Efrain Bleiberg, Anthony Bateman, and Liz Allison are listed (p. 544).

⁹ Fonagy, P. (2000). ‘Attachment and Borderline Personality Disorder’. *Journal of the American Psychoanalytic Association*, 48(4): 1129–1146, p. 1129.

¹⁰ Sandler, J. (1962). ‘Research in Psycho-Analysis—The Hampstead Index as an Instrument of Psycho-Analytic Research’. *The International Journal of Psychoanalysis*, 43: 287–291, p. 289. For varying examples of this scrutiny of psychoanalytic concepts, see e.g. Sandler, J. (1960). ‘On the Concept of Superego’. *The Psychoanalytic Study of the Child*, 15(1): 128–162; Sandler, J. and Rosenblatt, B. (1962). ‘The Concept of the Representational World’. *The Psychoanalytic Study of the Child*, 17(1): 128–145; Sandler, J., Dreher, A. U., and Drews, S. (1991). ‘An Approach to Conceptual Research in Psychoanalysis Illustrated by a Consideration of Psychic Trauma’. *International Review of Psycho-Analysis*, 18(2): 133–141. Sandler’s critical examination of the concept of the ‘preconscious’ can be regarded, in retrospect, as important groundwork for Fonagy’s work. See Sandler, J. and Sandler, A. M. (1994). ‘The Past Unconscious and the Present Unconscious: A Contribution to a Technical Frame of Reference’. *The Psychoanalytic Study of the Child*, 49(1): 278–292, p. 280.

¹¹ Sandler, J. (1983). ‘Reflections on Some Relations between Psychoanalytic Concepts and Psychoanalytic Practice’. *The International Journal of Psychoanalysis*, 64: 35–45, pp. 35–36.

The fact that they work well may in turn lead to an undue resistance to the progressive integration and modification of our concepts, so necessary for scientific development. This resistance can partly be overcome by the cultivation of a critical attitude towards our ideas, by discussion with colleagues, and by honest reading of the literature, but unless we are directly confronted with contradictions in our thinking, resistance to change due to secondary gains may prove too great.¹²

Sandler discussed various forms of ‘secondary gain’: use of imprecise or contradictory concepts can help avoid battles with authorities with stakes in the use of these concepts;¹³ it can help build apparent consensus; it allows a variety of people to project their own preconscious part-theories and fantasies on to the concept and feel satisfied with it. There may also be ways in which general concepts have their own acuity: after all, to see a blurred picture clearly is to see a blurred picture.¹⁴ Many of these forms of secondary gain, both positive and negative, can be seen in the case of Melanie Klein’s concept of ‘projective identification’ (see Chapter 6).¹⁵ However, Sandler felt that these forms of secondary gain should not distract from the harm unscrutinized concepts make to the effectiveness of causal claims, or the effectiveness of communication between groups.¹⁶

Fonagy has been a strong advocate of such attempts to confront contradictions and achieve integration in psychological theory: ‘while the clarification of terms and concepts is laborious, it is possible. It is also essential if we are to find out where theoretical differences are real, and to test these against each other, and where they may only be imagined.’¹⁷ Though these were certainly not rigid steps, in Fonagy’s view, the essence of Sandler’s approach was ‘to explore the history of a term or concept in psychoanalysis, then to elaborate on the multiple and frequently mutually incompatible meanings attached to the term. Having analysed the historical changes in terminology, he explains how misconceptions emerged or discussions at different levels of abstraction were conflated. Then, with a minimal number of assumptions, he proposes a highly economical model that encompasses the multiple uses of the varying meanings of the construct under scrutiny.’¹⁸ The function of such work is similar to a conceptual meta-analysis, as a concept’s aggregate relationships with other concepts are described and moderators identified that can account for variation between uses. Fonagy

¹² Sandler, J. (1962). ‘Research in Psycho-Analysis—The Hampstead Index as an Instrument of Psycho-Analytic Research’. *The International Journal of Psychoanalysis*, 43: 287–291, p. 290.

¹³ For further reflections on gerontocracy and the ensuing problems for conceptual development, see Fonagy, P. (2009). ‘When Analysts Need to Retire: The Taboo of Ageing in Psychoanalysis’, in B. Willock, R. Curtis, and L. Bohm (eds), *Taboo or not Taboo? Forbidden Thoughts, Forbidden Acts in Psychoanalysis and Psychotherapy*, London: Karnac Books, pp. 209–227.

¹⁴ Sandler, J. (1960). ‘On the Concept of Superego’. *The Psychoanalytic Study of the Child*, 15(1): 128–162: ‘the Index has functioned rather like a microscope, and as in the examination of physical tissues, increasing magnification may cause grosser structures to disappear from sight—but this by no means implies that they cease to exist’ (pp. 144–145). See also Wittgenstein, L. (1953). *Philosophical Investigations*, Oxford: Blackwell.

¹⁵ Sandler, J. (1987). ‘The Concept of Projective Identification’. *Bulletin of the Anna Freud Centre*, 10(1): 33–49. See also Sandler, J. (1993). ‘On Communication from Patient to Analyst: Not Everything is Projective Identification’. *The International Journal of Psychoanalysis*, 74: 1097–1107.

¹⁶ Sandler, J. (1987). *From Safety to Superego*, London: Karnac Books.

¹⁷ Fonagy, P. (2000). ‘On the Relationship of Experimental Psychology and Psychoanalysis: Commentary by Peter Fonagy (London)’. *Neuropsychanalysis*, 2(2): 222–232, p. 226.

¹⁸ Fonagy, P. (2005). ‘An overview of Joseph Sandler’s Key Contributions to Theoretical and Clinical Psychoanalysis’. *Psychoanalytic Inquiry*, 25(2): 120–147, p. 139. See also Bohleber, W., Fonagy, P., Jiménez, J. P., Scarfone, D., Varvin, S., and Zysman, S. (2013). ‘Towards a Better Use of Psychoanalytic Concepts: A Model Illustrated Using the Concept of Enactment’. *The International Journal of Psychoanalysis*, 94(3): 501–530. While the criteria proposed by Bohleber, Fonagy, and colleagues have not been followed strictly in examining the concepts of mentalization theory, they have been held in mind: relevance, falsifiability, operational definition, internal consistency, contextual consistency, parsimony, and (optional) extra-psychoanalytic convergence.

appreciated that Sandler's methodology encompassed both attention to the wider context and developments over time, and an attempt to proactively and constructively synthesize these developments to draw relevant distinctions in how a concept has been used. In this way, it allowed theory to 'stop and rewind' to identify the points at which different perspectives in use of a concept came to diverge and contribute to misunderstanding. In their 2003 book *Psychoanalytic Theories*, clearly influenced by Sandler, Fonagy and Target presented a historical and conceptual review of psychoanalysis attentive to the role of varying meaning of psychoanalytic concepts, and attempting to clarify and integrate difficulties these had caused.¹⁹ This included, impressively, a frank appraisal of limitations with their own theory and terminology at the time, and especially with the term 'mentalizing' (see Chapter 4).

With Sandler's studies in mind, as well as approaches from the sociology of science,²⁰ this book will explore 18 concepts that integrally organize the contributions of Fonagy and colleagues: adaptation, aggression, the alien self, culture, disorganized attachment, epistemic trust, hypermentalizing, reflective function, the p-factor, pretend mode, the primary unconscious, psychic equivalence, mentalizing, mentalization-based therapy, non-mentalizing, the self, sexuality, and teleological mode. As with Sandler's work, the analysis of key concepts is appreciative as well as critical, and intended to facilitate the success of the overall enterprise. We are mindful that matters such as terminological precision and the articulation of categories are not the priority of individual pragmatic researchers or clinicians. Furthermore, the work of Fonagy and colleagues offers an unprecedented integration of ideas from across different disciplines; it is inevitable that there remain some loose threads. Nonetheless, problems with the articulation of concepts and loose threads, which may be only minor irritants on a particular occasion, may cause wide-ranging issues for a field, played out incrementally over decades, with costs mounting.²¹ For instance, while the underspecification of concepts may contribute at times to their appeal by offering a screen for diverse projections, this predicament can also contribute to miscommunication, self-doubt, and reduced acuity when the concepts need to be used for practical work.²²

We have attempted to identify areas for theoretical refinement and loose threads, and, where possible, suggest a potential resolution or propose what future research needs to be done. For instance, in Chapter 4, we scope the diversity of ways the concept of 'mentalization' has been used—not as a criticism on the grounds of incoherence—but precisely as a means of attempting to synthesize a new definition. It is hoped that our overview and analysis will help those using the ideas of Fonagy and colleagues understand their underlying coherence and architecture, as well as a number of current limitations, and help researchers on mentalization and epistemic trust identify priorities for future research and theoretical development.

Our aim has been to offer reflections that are friendly and constructive, and actionable whenever possible. Exploration of tensions in the use of these concepts by Fonagy and colleagues is by no means a rejection of them, but rather an attempt to understand what problem

¹⁹ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications.

²⁰ For further reflections on method in our studies of psychological theory, see Duschinsky, R. (2019). 'Attachment and the Archive: Barriers and Facilitators to the Use of Historical Sociology as Complementary Developmental Science'. *Science in Context*, 32(3): 309–326.

²¹ This is a major theme of Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press.

²² See e.g. Hutsebaut, J., Bales, D. L., Busschbach, J. J., and Verheul, R. (2012). 'The Implementation of Mentalization-Based Treatment for Adolescents: A Case Study from an Organizational, Team and Therapist Perspective'. *International Journal of Mental Health Systems*, 6(1): 10.

or problems they aim to solve, how, and with what consequences for the theory as a whole. As Brecht observed in one of his dialogues, understanding a system of thought has analogies to understanding the dynamics of a family.²³ Much like any family, concepts in such a system can sometimes be seen supporting one another, sometimes changing and improving themselves to offer better support in the future; sometimes they can be seen squabbling, sometimes sharpening their knives for the next squabble. Like members of a family, each has to compromise their own integrity at times for the sake of the group: they may be forced to be more pragmatic, more complicated, more confused, or isolated from others outside the family, or even compromise links with reality for the sake of the family. At the same time, each member may also be sustained, in a deeper sense, by the family's collective capacities. The mode of life of concepts can be missed, however, if their self-presentation is taken at face value. It is not hard for concepts to sit down for dinner, or pose for a family photo, as if nothing had happened. But this may be far from the whole story.

The analysis of concepts entails careful thought about the language used to hold them. The language of psychological research and theory is a deceptively complex, multi-level system; this system has various entries and exits and opportunities for getting lost.²⁴ The complexity is further compounded by the interdisciplinarity of the work of Fonagy and colleagues, as terms from various discourses are taken up, often adapted, and worked into new configurations, retaining some of their previous connotations even as they become invested with a degree of technical or context-specific meaning. Fonagy and colleagues have attempted to avoid unnecessary abstraction, and remain close to ordinary language whenever possible. Their concepts have been developed not solely for academics or as a 'common language' for clinicians.²⁵ Another intended function has been for use in psychoeducation with patients. However, this then raises the ongoing challenge, in making sense of their work, of distinguishing between technical and ordinary uses of familiar words such as 'self', 'mental', 'adaption', and 'disorganization'. Careful scrutiny is therefore needed of the use of concepts across texts and over time to understand their meanings. For the book, definitions and uses of terms were identified from a complete examination of published works co-authored or authored by Peter Fonagy, Liz Allison, Anthony Bateman, Chloe Campbell, Marco Chiesa, György Gergely, Patrick Luyten, Howard Steele, Miriam Steele, and Mary Target. This was supplemented by study of relevant grey literature by these researchers. The writings of other researchers in the wider collaborative network—dubbed wryly by Fonagy and others as the 'mentalization mafia'²⁶—were also extensively consulted. These included Jon Allen, Eia Asen, Dickon Bevington, Efrain Bleiberg, Jessie Borelli, Martin Debbané, Karin Ensink, Jeremy Holmes, Elliot Jurist, Sigmund Karterud, Alessandra Lemma, Linda Mayes, Nick Midgley, Tobias Nolte, Carla Sharp, Finn Skårderud, Arietta Slade, and Lane Strathearn.

²³ Brecht, B. ([1961] 2019). *Refugee Dialogues*, trans. Tom Kuhn, London: Bloomsbury, p. 63.

²⁴ Fonagy, P. (2007). 'Interview', in L. E. Rubinstein (ed.), *Talking about Supervision: 10 Questions, 10 Analysts = 100 Answers*, London: International Psychoanalytic Association, pp. 39–49: 'Language carries its own intelligence. It's important to bear in mind that people's thinking is organized by their language. Knowing that helps you. Sometimes you make terrible mistakes: the language drives particular parts of content and you cannot drive it away from that' (pp. 43–44). See also Cavell, S. (1994). *In Quest of the Ordinary: Lines of Skepticism and Romanticism*, Chicago: University of Chicago Press.

²⁵ Fonagy, P. and Allison, E. (2014). 'The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship'. *Psychotherapy*, 51(3): 372–380, p. 375.

²⁶ E.g. Fonagy, P. (2013). 'Mentalization Based Interventions and a Mechanism of Change in Psychological Therapy'. *Improving Access to Psychological Therapies*. Accessed at: <http://www.uco.es/informacion/webs/fundacioncastilla/documentos/archivos/simposium/2013-simposio/presentaciones/fonagy-pres.pdf>.

Central collaborators

Fonagy has described opportunities for collaborative work as the greatest pleasure in his professional life.²⁷ Over time, he has co-authored works with countless colleagues. However, five sets of central collaborators can be picked out as playing a special role in pooling ideas and efforts.

A first set characterizes Fonagy's early work, during the 1980s. These are George Moran and Anna Higgitt.

George Moran served as Director of the Anna Freud Centre from October 1987 until his death at the age of 40 in January 1992. During these years, Fonagy was a trainee child analyst receiving supervision at the Centre. In turn, Moran was pursuing a PhD under Fonagy's supervision, exploring the relationship between blood sugar regulation and regulation of emotional state in diabetic children. Moran's doctoral research with Fonagy 'became the playing-working field on which a deep and mutually affectionate friendship and working relationship was established and flourished. It was a relationship in which mutual respect was enhanced by hours of discussion, debate and disagreement to force into view the best approximation of truth that friends and collaborators could muster.'²⁸

Anna Higgitt is a consultant psychiatrist and clinical leader in a community mental health team. From 1997, she has also held a role as senior policy adviser to the Department of Health.²⁹ Higgitt and Fonagy married in 1990.

A second set characterizes Fonagy's work in the early and mid-1990s. As well as Anna Higgitt, the two other key collaborators were Howard and Miriam Steele.

Howard and Miriam Steele arrived in London in 1986 to undertake doctoral research at University College London and clinical training at the Anna Freud Centre. Between 1987 and 1989, the Steeles conducted the Adult Attachment Interview with 100 expectant mothers and fathers in the third trimester. They also collected 96 infant–mother Strange Situation observations when the children were 12 months old, and 90 infant–father Strange Situation observations at 18 months. This was the basis for the University College London Parent–Child Project. Together with Fonagy, Howard and Miriam Steele, and Anna Higgitt developed the Reflecting Functioning Scale for the Adult Attachment Interview. Howard and Miriam Steele left London for the New School of Social Research in New York in 2001. They now hold Chairs in the Clinical Psychology Faculty and co-direct the Center for Attachment Research.

A third set are researchers with whom Fonagy elaborated the theory of mentalization in the 1990s and 2000s. Though diverse researchers were part of a wider community developing and thinking about these themes, Fonagy was especially influenced by his collaboration with Mary Target and György Gergely.

Mary Target pursued doctoral study with Fonagy, using a retrospective study of patient records at the Anna Freud Centre to evaluate the effectiveness of the service. Target had come from a background of a decade of work as a clinical psychologist in acute adult psychiatric services and child and adolescent mental health. She also trained as a psychoanalyst while completing her doctorate. Together with Fonagy, she published the influential series

²⁷ Fonagy, P. (2007). 'E-interview with Dominic Fannon'. *Psychiatric Bulletin*, 31(9): 360: 'What part of your work gives you the most satisfaction? Collaborating with colleagues in creating innovative treatment approaches, designing joint projects and writing as a team' (p. 360).

²⁸ Solnit, A. J. (1992). 'George Stritch Moran: A Personal Appreciation'. *Bulletin of the Anna Freud Centre*, 15(4): 267–268, p. 268.

²⁹ Baruch, G., Fonagy, P., and Robins, D. (eds) (2007). *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, John Wiley & Sons, p. x.

of papers on 'Playing with Reality' in the *International Journal of Psychoanalysis*, and a co-authored book, *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, in 2003. She is now Professor of Psychoanalysis at University College London, and a psychoanalyst in private practice. (In her clinical work, she practises under her maiden name, Mary Hepworth.)

György Gergely is a Hungarian clinical psychologist and experimental researcher associated with the Central European University in Budapest. Between 1996 and 1999, he was Visiting Senior Lecturer at University College London.³⁰ During this time, his concept of 'teleological mode' as a form of non-mentalizing was integrated into Fonagy's work. Together with Fonagy, Target, and Elliot Jurist, he was one of the authors of the 2002 book, *Affect Regulation, Mentalization and the Development of the Self*. He is also one of the originators of natural pedagogy theory, which helped prompt Fonagy's turn to concern with epistemic trust.

A fourth set of key collaborators are two researchers central to the development and validation of mentalization-based therapy (MBT), Anthony Bateman and Marco Chiesa. Increased funding was available in England for research on mental health services from 2006, thanks to the creation of the National Institute of Health Research. The growing reputation of the Anna Freud Centre also helped attract philanthropic donations to support research to develop and validate MBT, such as from the Borderline Personality Disorder Research Foundation and the Laurence Misener Charitable Trust.

Anthony Bateman is a consultant psychiatrist, Director of Psychotherapy Services and Research Lead at the St Ann's Hospital, London. He is also a practising psychoanalyst. Bateman and Fonagy developed MBT as a manualized treatment modality. Bateman's most recent publication is the second edition of the *Handbook of Mentalizing in Mental Health Practice*, co-edited with Fonagy.

Marco Chiesa was a consultant psychiatrist at the Cassel Hospital in Richmond from 1991 to 2015, where he served as Director of the outpatient Adult Personality Disorder Service. He also pursued a private practice as a psychoanalyst.³¹ Together with Fonagy, he was principal investigator of the Cassel Personality Disorder Study, a trial of a mentalization-based approach to therapeutic intervention with patients with personality disorders. These patients were also followed up over subsequent decades to examine the long-term implications of treatment. Subsequently, Chiesa led a study of patient characteristics and treatment pathways using data from 14 psychotherapy services. He has also collaborated with Fonagy in studying associations between personality disorder and reflective function.

A fifth set of significant collaborators have had an especially influential role in Fonagy's most recent work, and are co-authors on the landmark 2017 papers entitled 'What we have changed our minds about'.³²

Patrick Luyten is a clinical psychologist and psychodynamic psychotherapist, with faculty positions at University College London and KU Leuven. After qualifying as a clinician, Luyten had conducted empirical work on depression and chronic fatigue syndrome

³⁰ Details accessed at: <https://people.ceu.edu/sites/people.ceu.hu/files/profile/attachment/1600/gergelycv2012mayincl-natural-pedagogy-incl-richard-epiggle.pdf>.

³¹ Details accessed at: <https://www.ucl.ac.uk/psychoanalysis/people/marco-chiesa>.

³² Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What we have Changed our Minds about: Part 1. Borderline Personality Disorder as a Limitation of Resilience'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 11; Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What we have Changed our Minds about: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9.

at the University of Leuven and had previously collaborated with Sidney Blatt at Yale. It was through Blatt that Luyten was introduced to Fonagy. Luyten is now Director of the PhD in Psychoanalysis programme and Course Director of the PhD programme in Evidence-Based Child and Adolescent Mental Health at University College London. He also leads a treatment service for patients with depression and functional somatic disorders at KU Leuven.³³ His recent research has included work on the effectiveness of clinical interventions, and the development of self-report measures of mentalization.

Elizabeth Allison is Director of the Psychoanalysis Unit at University College London, and a practising psychoanalyst. She has a doctorate in English literature from Oxford University.³⁴ In writings with Fonagy, Allison has elaborated the idea of the primary unconscious, and the implications of epistemic trust for clinical practice.

Chloe Campbell is Deputy Director of the Psychoanalysis Unit at University College London. Her doctoral study addressed themes of race and empire in the history of Kenya.³⁵ Together with Fonagy, she is one of the authors of the article 'Bad Blood: 15 years on', addressing the relationship between attachment theory and mentalization approaches.³⁶ She has also contributed to current thinking about epistemic trust, especially drawing on interdisciplinary perspectives.

Overview of the book

After some biographical and contextual background in Chapters 1 and 2, Chapter 3 will detail Fonagy and Target's development of the idea of 'mentalizing' in thinking about the meaning of borderline personality disorder (BPD). BPD is a psychiatric diagnosis identified on the basis of problems with 'emotion dysregulation, impulsivity and social dysfunction'.³⁷ Chapter 3 will also consider the role of Howard and Miriam Steele and of attachment theory in shaping Fonagy's thinking about mentalizing and child development. Chapter 4 will close by outlining Fonagy's current position on mentalizing and identify transitions that have occurred over time. Chapter 5 will address the three modes of 'non-mentalizing' identified by Fonagy and Target. These will each be explained in turn. The chapter will also inspect suggestions in the work of Fonagy and colleagues regarding reasons why, when mentalizing breaks down, it falls into these three forms.

Having situated the ideas of mentalizing and non-mentalizing, we will then consider what these mean for understanding human psychology. Chapter 6 will critically examine a key term in Fonagy's vocabulary—'the self'—and consider the role of unconscious processes in the formation of thoughts and feelings. The chapter will also give attention to Fonagy's discussions of the 'alien self', and the role this psychological agency plays in sexuality and

³³ Details accessed at: <https://www.ucl.ac.uk/psychoanalysis/people/patrick-luyten>.

³⁴ Details accessed at: <https://www.ucl.ac.uk/psychoanalysis/people/liz-allison>.

³⁵ Campbell, C. (2010). 'Eugenics in Colonial Kenya', in A. Bashford and P. Levine (eds), *The Oxford Handbook of the History of Eugenics*, Oxford: Oxford University Press, pp. 289–300.

³⁶ Fonagy, P. and Campbell, C. (2015). 'Bad Blood Revisited: Attachment and Psychoanalysis, 2015'. *British Journal of Psychotherapy*, 31(2): 229–250.

³⁷ *The Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* lists nine criteria for BPD, of which five or more need to be met for a patient to receive a diagnosis. These are: unstable, intense relationships; inappropriate anger; frantic attempts to avoid abandonment; affective instability; impulsive and potentially self-damaging actions; recurrent self-injury or suicidality; chronic feelings of emptiness; paranoid thoughts or dissociative symptoms; and identity disturbances. See Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Borderline Personality Disorder', in Anthony Bateman and Peter Fonagy (eds) *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 323–334, p. 324.

in aggressive behaviour. Chapter 7 will explore the conceptualization of mental illness proposed by Fonagy and colleagues. The chapter will begin by considering the interest in adaptation among Fonagy and colleagues, as part of a broader trend within developmental psychology. The importance placed on learning from experiences with others in the origins of mental health and illness will be described. The chapter will also consider the critique of diagnostic categories offered by Fonagy and colleagues, and their account of the nature and structure of mental illness.

Chapter 8 will draw out the implications of the previous five chapters for how Fonagy and colleagues have approached the task of therapeutic intervention. The chapter will describe “Mentalisation-based Therapy,” and the evidence so far regarding its effectiveness. Chapter 9 will also attempt to synthesize the diverse reflections of Fonagy and colleagues on mentalization and non-mentalization in social systems, which they have stated will be ‘at the core of mentalizing endeavours in the future’.³⁸ The chapter will give attention to the work of Fonagy and colleagues on the school as an institution, and the potential for schools to contribute to the psychological well-being of their pupils. The chapter will also consider the researchers’ reflections on preventative interventions and public health infrastructures. Finally, the chapter will consider Fonagy’s reflections on the contribution of wider culture to the capacity of individuals and institutions to sustain mentalization and promote mental health and well-being. The Conclusion will draw together the claims from across the previous chapters. It will highlight the particular strengths of the work of Fonagy and colleagues, and identify a number of outstanding questions that face their paradigm.

³⁸ Bateman, A. and Fonagy, P. (2019). ‘Introduction’, in Anthony Bateman and Peter Fonagy (eds) *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, p. 18.

1

Biographical context

Introduction

On one level, this book has been designed to present an analysis of certain key concepts in the work of Fonagy and colleagues. However, simultaneously, we have also sought a structure that proceeds relatively chronologically, showing how later concepts built from earlier developments. One important such early development is the biographical context of Fonagy's concerns as a researcher. Even if collective processes such as conceptual discourses have a certain autonomy (see Chapter 9), they are always closely related to individual minds, and this shapes their meaning. Fonagy has himself urged awareness that 'behind the shifts in theories, techniques, and application are individuals whose sense of identity and ways of viewing the world have profoundly altered.'¹ Especially over the past decade, he has repeatedly emphasized the importance of the biographical context within which his ideas and collaborations emerged. This chapter will offer three biographical snapshots, selected for their relevance for understanding the later trajectory of his thinking, as well as on the basis of the available textual record. The first will be Peter Fonagy's experiences in the Fónagy household as a child in the 1950s. The second will be his psychoanalysis as a young man with Anne Hurry in the 1970s. Finally, close attention will be paid to Fonagy's work with one of his early patients.

The Fónagy family

Peter was born to Judith and Ivan Fónagy in Hungary on 14 August 1952. Peter Fonagy has described his father, Ivan Fónagy, as the single person who has most inspired him.² Ivan Fónagy was a Hungarian linguist, philosopher, and polymath. To develop his understanding of the development of language, he kept a diary of the early speech of his two children. This diary, excerpted in a book on Hungarian linguistics published in 1982, presents an initial view of the young Peter Fonagy.³ It also offers a sense of the intellectual preoccupations of the Fónagy home with both the developmental and interpersonal construction of meaning, and the developmental and interpersonal construal of difficult feelings. The chapter discussing the diary reflects Ivan Fónagy's characteristic concern to weave together observed particulars with general psychological phenomena. Psychoanalytic theory forms a general and

¹ Fonagy, P. (2000). 'Foreword', in Peter Fonagy, Robert Michels, and Joseph Sandler (eds), *Changing Ideas in a Changing World: The Revolution in Psychoanalysis. Essays in Honour of Arnold Cooper*, London: Karnac Books, pp. xvii–xxiv, p. vii.

² Fonagy, P. (2015). 'I Would Like to Abolish Silo Working'. *The Psychologist*, 28: 948. Accessed at: <https://thepsychologist.bps.org.uk/i-would-abolish-silo-working>.

³ Fónagy, I. (1982). 'He is Only Joking (Joke, Metaphor and Language Development)', in F. Kiefer (ed.), *Hungarian Linguistics*, Amsterdam: John Benjaminspp. 31–108.

subtly handled background to Ivan's interpretations of material from the diary, but without the theory defining his answers in advance.

The first, and perhaps the most prominent, use of young Peter's remarks in Ivan's reflections on linguistics was in discussion of how the concrete and metaphoric meanings of terms may bleed into one another for children. For instance, the word 'sweet' can be used to refer to a concrete experience of taste, or to a general experience of something good. Ivan observed Peter, aged 2 years and 9 months, extending from the concrete meaning to the metaphoric meaning in asking about the fate of unpopular books in terms of their lack of 'sweetness' to children: 'Books are printed again, says Peter's mother, if children like them.—And if they are not sweet? asks Peter.'⁴ Ivan reflected that such closeness of concrete phenomena and their associated potential meanings for children can make internal experience and external perception uncomfortably close and difficult to distinguish: 'There are no sharp borders between the personal, subjective and real universe which surrounds the child.'⁵

Second, Peter's remarks are used to illustrate the role of jokes and pretending in negotiating the terms under which an individual is inserted into meaning. Peter, aged 3, also told Ivan: 'Daddy is not allowed to eat my pudding.' When reproved by his grandfather, Peter retorted: 'I was just kidding.' Ivan was interested by the role that 'pretend' could play in such instances in turning an interaction into a 'nonreal activity'.⁶ Pretend here had the social function of extracting, or attempting to extract, the real force of a person's investment in and exposure to a social interaction. In Ivan's analysis, to be just kidding disowns, in this case strategically, the intentionality of wanting all the pudding and depriving father of access to it.

Finally, Peter's remarks are also used by Ivan to illustrate the role of caregivers in offering children access to the sense of what truth feels like. Ivan reported an interaction in which Peter, aged 3 years and 6 months, said to Judith: 'Mummy, I can see in your eyes that I was lying again.'⁷ Peter's statement was interpreted by Ivan as reflecting a condensation of the idea that the mother had 'seen through' Peter's lie. However, it was also interpreted as representing Peter's access to his own mind through the availability of the mother as a 'mirror', in which the lie could be reflected. For Ivan, the mother's face served as a mirror, offering Peter access to the perceptions of the other and, through this, access to his own mind over time.

There is a 20-year gap between the composition of the diary and Ivan's academic work in 1982 reporting from and reflecting on the diary. As a result, Ivan could therefore add the perspective of Peter, now aged '26 years and 8 months'. He reported that his son disapproved of his tendency to characterize the differences between childish and adult meanings as an effect of a child's shortcomings and confusion. This offered inadequate characterization of the psychological processes entailed:

Although I refer to these shortcomings of the child using the word confusion, clearly, this term does not accurately portray the complexity of the processes which must be involved in these errors. This thorny problem which needs further elaboration emerged in a discussion with Peter.⁸

⁴ Ibid. 56: 'The child will be even less willing to keep apart the shades of meaning of the same word. The semantic distance between original, primary "concrete" meaning and the derived, secondary "abstract" meaning is greatly reduced in the child's mind if not totally obliterated. The child is generally unwilling to renounce the concrete sensual interpretation of a word or idiom.'

⁵ Ibid. 75.

⁶ Ibid. 63.

⁷ Ibid.

⁸ Ibid. 70.

In 1992, Ivan and Peter presented a paper together on these questions at the International Psychoanalytic Association Conference on Psychoanalysis and Literature, a work published in 1995 in *Psychoanalysis and Contemporary Thought*. There, Fónagy and Fonagy argued that adult language has two components. One is the symbolic–linguistic code of differentiated adult meanings. The other is an ‘archaic accompaniment to language’, embedded in the *manner* in which something is said.⁹ The manner in which we do or say things easily escapes conscious control and is difficult to self-censor; it may therefore reflect aspects of childhood experience including embodied, associative thought. Fónagy and Fonagy proposed that cognitive impressions and communications contain within them vestiges of sensations and predispositions, which—deliberately or inadvertently—shape their tone. To dismiss something as an adult, for instance, may retain the vestiges of physical experiences of turning away, and the occasions through development in which a child’s circumstances led them to turn away.¹⁰ In turn, the cultural meanings ascribed to turning away will partially and dynamically shape its associated secondary senses.¹¹

Generally, the interplay between primary and secondary meanings is what enlivens speech, giving it depth. It permits a speaker to respond in their reply to the tone and manner in which something was said, both echoing aspects of the tone while also adding to or changing it to move the conversation forward.¹² A host of familiar, well-worn, social responses depend entirely on the concerted interplay of primary and secondary meanings, between what is said outright and what is implicated. For instance, in disinterested responses, the matter in hand is acknowledged but treated as without personal consequence.¹³ The concerted interplay of primary and secondary meanings permits discretion, good-humouredness, irony, gentle teasing, tentativeness, low-level threats, implicit requests, and many other kinds of reaction in which a meaning is at once offered and qualified.¹⁴ Consider the variety of ways in which routine phrases like ‘have a nice day’ can be inflected. Or again, note how sounds (e.g. ‘p’, ‘t’, ‘k’) associated with lightness and fastness may be used in brand names that seek to convey these properties to consumers, for instance medications for chemotherapy.¹⁵ However, it is also possible for the two components of adult language to run contrary to one another, with one component of language closing down or tripping up the other:

⁹ Fónagy, I. and Fonagy, P. (1995). ‘Communication with Pretend Actions in Language, Literature and Psychoanalysis’. *Psychoanalysis and Contemporary Thought*, 18(3): 363–418, p. 366.

¹⁰ The example comes from an exposition of the Fónagy and Fonagy paper in Fonagy, P. and Target, M. (2007). ‘The Rooting of the Mind in the Body: New Links between Attachment Theory and Psychoanalytic Thought’. *Journal of the American Psychoanalytic Association*, 55(2): 411–456, p. 433.

¹¹ This was discussed at more length some years later, in Fonagy, P. (2012). ‘Does it Matter if There Is a Nonverbal Period of Development? On the Infant’s Understanding the Social World and its Implications for Psychoanalytic Therapy’. *Journal of the American Psychoanalytic Association*, 60(2): 287–296.

¹² This idea is already present in Ivan and Judith’s paper on the melodic patterns of speech: Fónagy, I. and Fónagy, J. (1987). ‘Analysis of Complex (Integrated) Melodic Patterns’, in R. Channon and L. Shockey (eds), *In Honour of Ilse Lehisté*, Dordrecht: The Netherlands: De Gruyter Mouton, pp. 75–98. ‘Foris: “Enumerative information patterns may appear in French sentences which do not contain enumeration conveying secondary messages such as ‘for the one thing’, there are many other reasons of course’”’ (p. 75).

¹³ See e.g. Maynard, D. W. and Hudak, P. L. (2008). ‘Small Talk, High Stakes: Interactional Disattentiveness in the Context of Prosocial Doctor–Patient Interaction’. *Language in Society*, 37(5): 661–688.

¹⁴ Fónagy, I. (1995). ‘Iconicity of Expressive Syntactic Transformations’, in M. E. Landsberg (ed.), *Syntactic Iconicity and Linguistic Freezes: The human dimension*, Berlin: Mouton de Gruyter, pp. 285–304; Galgut, E. (2010). ‘Reading Minds: Mentalization, Irony and Literary Engagement’. *The International Journal of Psychoanalysis*, 91(4): 915–935; Culpeper, J., Haugh, M., and Sinkeviciute, V. (2017). ‘(Im)politeness and Mixed Messages’, in J. Culpeper, M. Haugh, and D. Z. Kádár (eds), *The Palgrave Handbook of Linguistic (Im)politeness*, London: Palgrave, pp. 323–355.

¹⁵ Abel, G. A. and Glinert, L. H. (2008). ‘Chemotherapy as language: Sound symbolism in cancer medication names’. *Social Science & Medicine*, 66(8): 1863–1869.

Some borderline individuals may appear to ‘ask’ questions whilst at the same time letting their analyst know that they are already aware of the answer. For example, one patient was in the habit of ‘asking’ relatively frequently: ‘Are you bored?’ As there was no inflection at the end of the sentence, the implication was that, whether through boredom or some other reason, the analyst was not doing his job properly.¹⁶

Peter was evidently proud of this paper written with his father.¹⁷ Whereas in the 1980s, his father had characterized children’s thought as confused, the collaboration with Peter had prompted a development in this position. Secondary meanings are present across human speech in both childhood and adulthood. Furthermore, they may be regarded as part of how sense and relevance are created and communicated. To claim to be just kidding after asserting exclusive rights over pudding may offer a way of defusing the aggressive aspects of the assertion. Removing inflection from the end of a question undercuts it as a question, forming it into an unstated and unanswerable accusation. The interplay of primary and secondary meanings, for children and for adults, signals an exposure to our history that may be qualified, stylized, or—at times—may ambush us. However, neither Ivan nor Peter at this point acknowledged, as would Peter many decades later, that exposure to history also means exposure to culture (see Chapter 9).¹⁸

In the 1990s, Fónagy and colleagues wrote versions of a paper published first in the *Bulletin of the Anna Freud Centre* and then in the *Journal of the American Psychoanalytic Association*. In this paper, the authors cited an anonymous ‘guilty 4-year-old who said to his mother “I can tell from your eyes that I am lying.”’¹⁹ In the first version of the paper, Fónagy and colleagues commented that such a statement does not signify a child who has regressed:

to a stage of ‘fusion’ with his primary object. Rather, it was his grasp of the boundaries of his mental world which appeared to be insufficiently well established. The initial reliance on the object’s reflective self and the gradual recognition of its separateness are both necessary to the evolution of a securely established reflective self.

In the second version of the paper, after arguing against an explanation in terms of regression, the authors proposed that ‘it is more likely that his grasp of the boundaries of his mental world momentarily returned to an earlier state (in this instance, consequent on anxiety and self-reproach).’²⁰

¹⁶ Fónagy, I. and Fónagy, P. (1995). ‘Communication with Pretend Actions in Language, Literature and Psychoanalysis’. *Psychoanalysis and Contemporary Thought*, 18(3): 363–418, pp. 385–386.

¹⁷ Bach, S., Alvarez, A., Mayes, L., and Fónagy, P. (2000). ‘Panel 3: Fantasy Life and the Self’. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 51–62. ‘It is the only paper I ever managed to write with my father, and Mary Target, who assisted the birth of this paper will, I think, vouch that it was a difficult and painful delivery. Dad, are you listening? I should tell you that my father is a linguist and a phonetician. He has, I believe, one of the most profound and psychoanalytically relevant theories of language development around’ (p. 58).

¹⁸ In the 1980s, Derrida criticized Ivan Fónagy’s work for neglecting adequate acknowledgement of cultural signification in assuming too unmediated a relationship between preverbal childhood experience and adult sense-making. Derrida, J. (1986). *Glas*, trans. Richard Rand and John P. Leavey, Lincoln, Nebraska: University of Nebraska Press.

¹⁹ Fónagy, P., Steele, M., Moran, G. S., Steele, H., and Higgitt, A. C. (1991). ‘Measuring the Ghost in the Nursery: A Summary of the Main Findings of the Anna Freud Centre/University College London Parent-Child Study’. *Bulletin of the Anna Freud Centre*, 14: 115–131, p. 127.

²⁰ Fónagy, P., Steele, M., Moran, G., Steele, H., and Higgitt, A. (1993). ‘Measuring the Ghost in the Nursery: An Empirical Study of the Relation between Parents’ Mental Representations of Childhood Experiences and Their Infants’ Security of Attachment’. *Journal of the American Psychoanalytic Association*, 41(4): 957–989, p. 983.

In 1982, Ivan had emphasized the role of the mother as a 'mirror' offering little Peter access to his own mind and experience of truth. In the 1990s, Fonagy did not exactly disagree. The lie did offer a vehicle for exploring internal reality through the other's response.²¹ At the same time, though, Fonagy's sense was that this was not solely a positive exploration and reflected something insufficiently well established. He recalled himself as a child with 'lots of anxieties and lots of difficulties',²² offering a different slant on the incident. Looking back over the decades, Fonagy suggested that little Peter's anxiety and self-reproach for his lie prompted a shift to a state in which he was unable to know the reason for his anxiety and self-reproach except through the concrete observable behaviour of his mother. The result of the lie is the mother's concrete, observable response, and it is this that makes the deception apparent, not little Peter's knowledge of his own mind. The papers from the 1990s show Fonagy's reflection on, and renegotiation of, his father's frame of reference. This included his father's concern with experience–reality confusion, the use of pretend, and the interpretation of intentions from observable behaviour. The papers also show the way that Fonagy was building from these concerns to think about the limitations we may face in understanding our own minds, especially in the context of attachment relationships. Part of the context of these reflections was a preoccupation with self-understanding and recognition of feeling states among some clinicians at the Anna Freud Centre in the 1970s, including Fonagy's own therapist, Anne Hurry.

Fonagy's psychoanalysis with Anne Hurry

In 1970, Ivan Fónagy escaped communist Hungary, took up a professorship in linguistics in Paris, and began training as a psychoanalyst. However, Ivan felt that Peter should have an English education, so he sent his son, then aged 14, on his own to live in London.²³ The young Peter knew no English and had no friends.²⁴ He came to dread getting up each morning to go to school, where he was bullied by peers.²⁵ He felt 'sad, constantly worried and inadequate',²⁶ and frequently had suicidal thoughts. He recalls that 'as a young person, one of us (PF) used to phone home in states of distress and talk about his situation in catastrophic terms until his parents were palpably panicked, and then he would end the conversation feeling relieved'.²⁷ By age 17, Peter was depressed, suicidal, and failing academically. A neighbour worked at the Hampstead Child Therapy Course and Clinic, and encouraged Fonagy to seek treatment.²⁸

²¹ See also Lemma, A. (2005). 'The Many Faces of Lying'. *The International Journal of Psychoanalysis*, 86(3): 737–753: 'Lying can be thought about as part of this process. Through the lie the child can transform something untrue/unreal into something "real" by placing a lie about the self inside the object's mind (Fonagy, personal communication). In other words, the lie provides a vehicle for exploring one's own internal reality through the object's response to it, thereby making it "real" for the self' (p. 748).

²² Ezrati, O. (2014). 'Freud Off: Giving New Meaning to Psychoanalysis'. *Haaretz*, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>.

²³ Fonagy, P. (2015). 'Peter Fonagy on Psychoanalysis and IAPT'. *The History of Emotions Blog*, posted on 14 May by Jules Evans. Accessed at: <https://emotionsblog.history.qmul.ac.uk/2015/05/peter-fonagy-on-psychoanalysis-and-iapt>.

²⁴ Fonagy, P. (2015). 'Peter Fonagy on Psychoanalysis and IAPT'. *The History of Emotions Blog*, posted on 14 May by Jules Evans. Accessed at: <https://emotionsblog.history.qmul.ac.uk/2015/05/peter-fonagy-on-psychoanalysis-and-iapt>.

²⁵ BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. *The Life Scientific Podcast*, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

²⁶ Fonagy, P. (2020). Foreword to Leo Potton and Ana Strumpf, *Perfectly Imperfect Stories*, London: Magic Cat Publishing.

²⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 21.

²⁸ Kirby, T. (2019). 'Peter Fonagy—Battling the Enemy of Loneliness'. *The Lancet Psychiatry*, 6(12): 987.

The initial assessment meeting did not go well, and he felt so uncomfortable that he resolved never to go back.²⁹ However, he reversed this decision six months later as his feelings of loneliness and depression became more entrenched.³⁰ He went into analysis with Anne Hurry for what would be three years of treatment. The Hampstead Clinic was renamed the 'Anna Freud Centre' in 1982, following the death of Anna Freud. (For convenience, we will henceforth refer to the Anna Freud Centre even prior to 1982.)

Hurry's therapeutic approach was an expression of important developments in clinical technique that had been germinating at the Anna Freud Centre over the previous decades. In the 1950s, Anna Freud circulated unpublished papers to her colleagues.³¹ There she suggested that some individuals may not be suitable straightaway for the 'classical technique' of psychoanalytic interpretations punctuating long silences. The patients first needed to receive support in establishing relationships within which the silences and the interpretations could be meaningful and useful. She proposed that many of these patients had experienced traumatic events in their relationships with their caregivers.³² In the 1960s, Freud expanded her stance. She argued that, in general, many forms of therapeutic work could benefit from this additional approach, which she labelled as 'developmental help'.³³ Developmental help was taken to mean that a therapist should offer a patient undivided attention, help the patient recognize and verbalize feelings, develop an account of themselves and their needs, and clarify the reality of situations and of cause and effect. In this, Freud pitted her clinical approach against that of Melanie Klein, whom she felt ignored children's maturational processes and the importance of present-day experiences, including the child's contemporary experiences of care by the parent.³⁴

Despite her advocacy for alterations in clinical technique, Anna Freud was somewhat ambivalent about developmental help, and the prospect of departing from the approach espoused in her father's writings.³⁵ Anna Freud's personal ambivalence no doubt contributed to a diversity of opinion at the Centre. Nonetheless, among many therapists associated with the Centre, the technique of developmental help was central to work with children and adolescents. It was anticipated that these interventions would help unblock pathways towards

²⁹ Fonagy, P. (2020). Foreword to Leo Potton and Ana Strumpf, *Perfectly Imperfect Stories*, London: Magic Cat Publishing.

³⁰ BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

³¹ Freud, A. ([1956] 1969). 'The Assessment of Borderline Cases', in *The Writings of Anna Freud*, Volume 5. New York: International Universities Press, pp. 301–314.

³² Ibid. 310.

³³ Freud, A. (1965). *Normality and Pathology in Childhood*, Harmondsworth, UK: Penguin Books. See also Freud, A. (1971). 'The Infantile Neurosis: Genetic and Dynamic Considerations'. *The Psychoanalytic Study of the Child*, 26(1): 79–90: 'The analyst's therapeutic ambition goes beyond the realm of conflict and the improvement of inadequate conflict solutions. It now embraces the basic faults, failures, defects and deprivations, i.e. the whole range of adverse external and internal factors, and it aims at the correction of their consequences' (p. 89).

³⁴ The divergences between Freud and Klein on clinical technique are discussed by Fonagy and colleagues in Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38. See also Midgley, N. (2012). *Reading Anna Freud*, London: Routledge.

³⁵ Edgumbe, R. (1995). 'The History of Anna Freud's Thinking on Developmental Disturbances'. *Bulletin of the Anna Freud Centre*, 18(1): 21–34: 'Over the years we have felt uncertainty about the use of "non-analytic techniques" ... Anna Freud made us feel that somehow it was not right. Yet she undoubtedly encouraged us to carry on experimenting' (p. 22). This allegiance to psychoanalytic interpretation as a symbolic marker of 'true' psychoanalysis, even while clinical practice has shifted in the direction of supportive counselling, has been discussed by Lecours, S. (2007). 'Supportive Interventions and Nonsymbolic Mental Functioning'. *The International Journal of Psychoanalysis*, 88(4): 895–915. The extent to which Sigmund Freud himself ever adopted 'classical' technique is unclear from the available evidence: Lohser, B. and Newton, P. M. (1996). *Unorthodox Freud: The View from the Couch*, New York: Guilford Press.

healthy development, and facilitate the therapeutic relationship between patient and therapist.³⁶ While developmental help was increasingly treated as applicable to all therapeutic work, in the 1960s, The Group for the Study of Borderline Cases pursued the question of so-called borderline states, affirming that such patients are, especially, in 'no condition to co-operate or fully comprehend interpretations. It seems, however, that the therapist's empathy comes across at such times.'³⁷ The boundaries of what constituted a 'borderline state' were amorphous.³⁸ Nonetheless, the Group identified that these patients were characterized especially by distress, difficulties in personal relationships, and an immediate and urgent translation of experiences into action, as a result of difficulties in holding on to thoughts.³⁹ Rosenfeld and Sprince also reported the impression of members of the Group that the interpretation of psychological conflicts with these patients tended to backfire.⁴⁰

By the 1970s, Hurry was aligned with this perspective, and had come to advocate the clinical use of empathy rather than interpretations alone.⁴¹ At the time that Fonagy entered into therapy with her, Hurry had recently joined a new study group focused on the role of patient self-understanding within clinical practice. Discussions led the group to conclude that patients would benefit not only from 'insight' into what had happened to them in the past, but also self-understanding regarding present-day experiences, and the role of these experiences in serving to impede or support their coping and development.⁴² To achieve

³⁶ Cf. Szydlo, J. S. (1985). 'Developmental Help: Intensive Treatment of a Nursery Child'. *Bulletin of the Anna Freud Centre*, 8(1): 23–38, p. 32; Fonagy, P. and Target, M. (1996). 'Predictors of Outcome in Child Psychoanalysis: A Retrospective Study of 763 Cases at the Anna Freud Centre'. *Journal of the American Psychoanalytic Association*, 44: 27–77. 'Developmental help includes those aspects of the analyst's interaction with the child which address: (1) ego functions (mental processes) via self and object representations; (2) the verbalization of internal states and differentiation of affects; (3) the breaking down of unmanageable affects (anxiety) into smaller manageable entities which the child can master; (4) the development of internal representations of affects so that the child can master his own feelings; (5) the facilitation of thinking by reducing anxiety and making links between different aspects of thought processes; (6) facilitation of thinking about cause and effect, particularly within relationships; (7) helping the child separate internal from external, real from unreal, fantasy from reality; (8) setting limits and offering explanations for the limits provided; (9) facilitating the creation of internal representations of self and other; (10) establishing reciprocity (e.g. giving and taking); (11) developing the capacity to delay gratification; (12) helping the child to develop an "as if" attitude, and the encouragement of fantasy; (13) gradually confronting the child with opposing ideas, for example, the possibility of hatred and dependence on the same person. While these elements enter into any analysis, especially with children, we suggest that they need to be much more intensively and repeatedly focused on with children whose primary disturbance is at the level of mental processes' (p. 61).

³⁷ Rosenfeld, S. K. and Sprince, M. P. (1965). 'Some Thoughts on the Technical Handling of Borderline Children'. *The Psychoanalytic Study of the Child*, 20(1): 495–517, p. 509. The specific case under discussion here was later considered again by Hepworth and Fonagy based on examination of the case record. Target, M. and Fonagy, P. (2003). 'Attachment Theory and Long-Term Psychoanalytic Outcome: Are Insecure Attachment Narratives Less Accurate?', in M. Leuzinger-Bohleber, A. U. Dreher, and J. Canestri (eds), *Pluralism and Unity? Methods of Research in Psychoanalysis*, London: International Psychoanalytical Association, pp. 149–167.

³⁸ New, A. S. and Triebwasser, J. (2018). 'A History of Borderline Personality Disorder', in Barbara Stanley and Antonia New (eds), *Borderline Personality Disorder*, Oxford: Oxford University Press, pp. 1–16. Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications: 'With the benefit of a historical perspective, we may readily see the group of disorders she describes as ultimately analogous to disorders of personality ... Her concept of non-conflictual disorders was revolutionary and the linking of these pathologies to development prescient. Her developmental perspective, however, was never properly reconciled with structural theory' (p. 77).

³⁹ Rosenfeld, S. K. and Sprince, M. P. (1963). 'An Attempt to Formulate the Meaning of the Concept "Borderline"'. *The Psychoanalytic Study of the Child*, 18(1): 603–635, p. 625. The initial members of the Group for the Study of Borderline Cases were Bene, Fahmy, Ini, Kawenoka, Kennedy, Rosenfeld, Schnurmann, Singer, and Sprince.

⁴⁰ Rosenfeld, S. K. and Sprince, M. P. (1965). 'Some Thoughts on the Technical Handling of Borderline Children'. *The Psychoanalytic Study of the Child*, 20(1): 495–517.

⁴¹ Hurry, A. (1977). 'Living under the Threat of Death: Discussion'. *Journal of Child Psychotherapy*, 4(3): 22–24: 'We all meet cases who need for periods to create their own worlds, and where interpretations are simply an impingement.'

⁴² Kennedy, H. (1979). 'The Role of Insight in Child Analysis: A Developmental Viewpoint'. *Journal of the American Psychoanalytic Association*, 27: 9–28. Fonagy would later emphasize his intellectual debt to this paper, and the study group discussions that led to its composition. Fonagy, P. and Moran, G. S. (1991). 'Understanding

this, Hurry felt it was important to 'give the adolescent some experience of being understood. This does not normally involve any "deep interpretation", but simply a recognition of his feeling state'.⁴³

Hurry's approach in analysis with suicidal and depressed adolescents in the 1970s is suggested by a 1977 paper in the *Journal of Child Psychotherapy*. There, she described her treatment of an adolescent girl, Jessie. The overriding theme of the therapeutic work was the girl's feeling of being unprotected by her mother. In this case, as in others, Hurry attempted to be kind and nurturing to her adolescent patient. However, she described her resoluteness in defending the distinction between reality and pretend, in the belief that reality is ultimately reassuring for all that it is also disappointing. Her impression was that a fantasy of omnipotence was often active for suicidal adolescents as a reaction against feelings of helplessness or self-recrimination. Though a defence against feelings of fear and guilt, the sense of omnipotence could also at times exacerbate these feelings and also predispose suicidal thoughts.⁴⁴ In the case of Jessie:

For a long time she tested out how far I resembled her mother in willingness to let her be exposed to danger. Once she asked me to lie for her: to write a certificate saying she was eighteen so that she could go to a very unsuitable club. The next day she told me triumphantly that mother had helped her to forge a certificate so that she could get in. But in fact she was very relieved that I had not lied.⁴⁵

Reflecting on his own therapy with Hurry, Fonagy recalled that her technique 'wasn't very classical'. One of Hurry's major concerns was supporting her patients to gain cognitive insight, using the therapeutic relationship as a focus. However, experiences outside the analytic situation were also acknowledged as having their own reality to deplete or support the patient, motivating kind and validating responses by the analyst:

One of the moments I remember from it is that, about a year into my analysis, I had my first car, which was an old banger—a Ford Anglia—and I went to lay down on the couch and said, Such a wonderful car, it's great. She [Hurry] stopped, said hold on a second, went to the window and looked out, came back, sat down, and said: Peter, it's a wonderful car. This I remember from my analysis. She gave space to reality. Exactly. She was a very, very human

Psychic Change in Child Psychoanalysis'. *The International Journal of Psychoanalysis*, 72: 15–22: 'Kennedy highlights the child analyst's focus on the patient's present conflicts and concerns. In this context, the analyst aims to foster progressive development and improved adaptation by providing the patient with a framework which will allow him to understand himself better' (p. 16). See also Target, M. (2010). 'The Psychoanalytic Work of Hansi Kennedy: From War Nurseries to the Anna Freud Centre (1940–1993)'. *Journal of Child Psychotherapy*, 36(3): 313–316, though other sources also pointed Fonagy in the same direction. See e.g. Fonagy's 1989 review of Harold F. Searles, M.D. (1986). *My Work with Borderline Patients*, Northvale, NJ and London: Jason Aronson, in *Bulletin of the Anna Freud Centre*, 12(1): 77–79.

⁴³ Hurry, A. (1986). 'Walk-in Work with Adolescents'. *Journal of Child Psychotherapy*, 12(1): 33–45, p. 41.

⁴⁴ Hurry was unpersuaded by the idea of a death drive. Hurry, A. (1978). "My Ambition is to be Dead": The Analysis of Motives and Reasons for Suicide Behaviour in an Adolescent Girl, with Particular Reference to the Relationship between the Adolescent Process and Suicide: Part II: Past and Current Findings on Suicide in Adolescence'. *Journal of Child Psychotherapy*, 4(4): 69–79: 'This latter theory has not been universally accepted, and I myself find it an impediment rather than an aid in the analysis of suicidal patients' (p. 75).

⁴⁵ Hurry, A. (1977). "My Ambition is to be Dead": The Analysis of Motives and Reasons for Suicide Behaviour in an Adolescent Girl, with Particular Reference to the Relationship between the Adolescent Process and Suicide: Part I: Case Study'. *Journal of Child Psychotherapy*, 4(3): 66–83, p. 71.

person, and what adolescents need—what I needed then—was really someone who was going to relate to me and see me where I was.⁴⁶

Fonagy has repeatedly stated that therapy with Hurry saved his life,⁴⁷ helping him to find meaning in 'the present day and the here-and-now'.⁴⁸ He has described how 'the power of personal recognition,' provided by Hurry, 'gave me courage to internalize and organize myself and my life in very many respects in her image of me'.⁴⁹

Fonagy has reflected that a lasting benefit of his work with Hurry was the refiguration of his depression into 'an enduring commitment to confronting unpleasant truths and understanding'.⁵⁰ Many of the concerns of his professional life can be regarded as, to a certain extent, reflecting questions he faced as an isolated adolescent émigré to Britain: attachment and separation (Chapter 4); challenges of social understanding and the opacity of other minds (Chapter 3); feelings of incoherence and meaninglessness, and of alienness (Chapter 6); the advantages and price of adaptation to the environment (Chapter 7); the need to learn from others (Chapter 7); and the complexities and potential cruelties of the school environment and of culture (Chapter 9). As a clinician, Fonagy has described that it has 'helped me enormously' to be able to 'reach back to that part of myself' and use reflection on the distress and hopelessness of his teenage years in considering the experience of his patient, though mindful also of the differences between his difficulties and those of the other.⁵¹

Despite its important successes, the therapeutic achievements with Hurry had limitations. As Fonagy and colleagues have subsequently argued, long-term therapy with a focus on supporting cognitive insight may have limitations for reducing non-mentalizing (Chapter 8). Fonagy would be forced to re-enter analysis again as a young adult. And throughout his life he would report struggling with persistent depressive disorder and destructive workaholicism. At times, depression and workaholicism would assuage, and at times exacerbate, one another.⁵² Fonagy has been able to pursue the development of theory drawing on relevant experiences as a clinician, as a researcher, and also as a service user. At a conference for users of psychological services:

The house lights stay on throughout, because this is not an 'expert' talking at 'service users', but a man who appears to care deeply about his work and to want to share a little of his

⁴⁶ Ezrati, O. (2014). 'Freud Off: Giving New Meaning to Psychoanalysis'. *Haaretz*, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>.

⁴⁷ Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 90.

⁴⁸ Fonagy, P. (2015). 'Peter Fonagy on Psychoanalysis and IAPT'. The History of Emotions Blog, posted on 14 May by Jules Evans. Accessed at: <https://emotionsblog.history.qmul.ac.uk/2015/05/peter-fonagy-on-psychoanalysis-and-iapt/>.

⁴⁹ Fonagy, P. and Allison, E. (2018). 'The Origin of Human Life: A Psychoanalytic Developmental Perspective', European Psychoanalytical Federation, 31st Annual Conference, Warsaw, 24 March.

⁵⁰ Fonagy, P. (2007). 'E-interview with Dominic Fannon'. *Psychiatric Bulletin*, 31(9): 360.

⁵¹ BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

⁵² Shragai, N. (2014). 'Business Travel and Families: The Ultimate Guilt Trip'. *Financial Times*, 13 October. Accessed at: <https://www.ft.com/content/f7eea390-4d66-11e4-bf60-00144feab7de>; Ratner, R. H. (2014). 'Interview with Peter Fonagy', in *Borderline*. Master of Fine Arts, The University of Texas at Austin. Accessed at: <https://repositories.lib.utexas.edu/bitstream/handle/2152/28667/RATNER-MASTERSREPORT-2014.pdf>, p. 123; Fonagy, P. (2016). 'Foreword', in Cooper, A. and Redfern, S. (eds). *Reflective Parenting: A Guide to Understanding What's Going on in Your Child*, London: Routledge, pp. ix–xii: 'It doesn't happen to me often that I feel worthwhile. Most days I do what I feel I have to do, and if I have done 50% of what I needed to, I feel good. The outcome I am for is just to have coped' (p. ix). Approval of this discussion of his struggles around mental health was given explicitly by Peter Fonagy (personal communication, December 2019).

knowledge with people who really need it. Any barriers between ‘them’ on the stage and ‘us’ in the audience are further broken down when Professor Fonagy shared with us his personal experience of major depression and suicidal thoughts in the past.⁵³

Fonagy holds as a personal value that mental health needs and use of mental health services should be discussed openly rather than treated as a source of shame. Doing so can help acknowledge both the limitations and the achievements of psychological therapies, and this has been the context for his frequent discussions of his own history in his writings and interviews especially over the past decade.⁵⁴ The achievements of Fonagy’s therapy with Hurry were clearly very extensive and helped sustain an unusual configuration: a deeply felt confidence in the capacity of psychoanalysis to help shed light on truths about the human mind and to help people who are suffering;⁵⁵ and an empiricist impatience with anything, including psychoanalysis, that seemed unwilling or unable to confront reality in the final instance. This configuration represented a transformation of depressive experiences, but also retained a continuity with them. Fonagy’s position on the ‘fault line’ between psychoanalysis and empirical psychology is one he has explicitly associated with a childhood fantasy of dread: ‘the gulf, gap, or chasm between psychoanalysis and psychology generated associations to dreams I used to have as a child where, standing on an elevated, unstable surface, I suddenly found myself falling and aware of the inevitability of a disastrous end.’⁵⁶

One quality Fonagy appreciated in Hurry as an analyst was that, alongside helping him to ‘say truths that I didn’t like and yet retain a sense of myself as worthwhile,’ Hurry was ‘very funny to boot.’⁵⁷ Hurry’s case reports describe making space for fun and humour.⁵⁸ On the surface, this might seem opposed to her emphasis on reality. Having fun may function as a form of ‘just kidding,’ in which an interpersonal or a personal reality is shrugged off. However, Fonagy appreciated Hurry’s humour and its contribution to their work together. Praising the importance of humour, he would later reflect that ‘rigid seriousness can exaggerate threats, amplify slights and close down our openness to different ways of responding to things. Humour can protect us; it gives us the ability to see a situation from another point of view, and gives perspective when we might feel overwhelmed.’⁵⁹ Humour, even outright pretence, offers a renewed and different access to reality through a step outside our own habitual stance, precisely through the utilization of an interplay of primary and secondary meanings. Humour can counteract the risk of overstating the equivalence between internal perception and external states of affairs. And they can counteract the risk that external action is overstated in its importance, at the expense of seeking perspective on it.

⁵³ See e.g. Walker, S. (2011). ‘It’s Good to Talk, 2: Where’s our Support Group?’ Accessed at: <https://purplepersuasion.wordpress.com/2011/09/14/its-good-to-talk-2-wheres-our-support-group/>.

⁵⁴ See e.g. BBC Radio 4 (2020). ‘Peter Fonagy on a Revolution in Mental Health Care’. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

⁵⁵ Ezrati, O. (2014). ‘Freud Off: Giving New Meaning to Psychoanalysis’. Haaretz, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>: ‘Psychoanalysis has more truth to it about the mind than any other psychology I know. My own experiences on the couch—and I spent about 18 years there—helped me immeasurably with understanding people.’

⁵⁶ Fonagy, P. (2000). ‘On the Relationship of Experimental Psychology and Psychoanalysis: Commentary by Peter Fonagy (London)’. *Neuropsychanalysis*, 2(2): 222–232, p. 222.

⁵⁷ Fonagy, P. (2015). ‘I Would Like to Abolish Silo Working’. *The Psychologist*, 28: 948. Accessed at: <https://thepsychologist.bps.org.uk/i-would-abolish-silo-working>.

⁵⁸ E.g. Hurry, A. (1998). ‘“Paul”: The Struggle to Restore a Development Gone Awry’, in A. Hurry (ed.), *Psychoanalysis and Developmental Therapy*, London, Karnac Books, pp. 100–123.

⁵⁹ Fonagy, P. (2016). ‘Laughter, Trauma and Building Resilience’. Huffington Post, 31 July. Accessed at: https://www.huffingtonpost.co.uk/peter-fonagy/comedy-mental-health_b_11262496.html.

True, humour may be deployed to deflect thoughts and feelings. Nonetheless, in general, 'humour can be considered to be a coping strategy and a positive indication of mentalizing abilities; it attests a person's ability to "play with ideas";⁶⁰ Fonagy and Bateman would explicitly incorporate humour into their guidebook for clinical interventions. They encouraged the use of 'mischievous statements', provocative and left-field, when a patient seems stuck in rigid seriousness. They likewise prescribed the use of 'wacky statements', puzzling and unexpected, when a dose of curiosity is needed.⁶¹ Fonagy and Bateman described the 'inquisitive stance' as a 'style of interaction characterized by an expectation that an individual's mind may be influenced, surprised, changed, and enlightened by learning about another's mind.'⁶² From his descriptions of analysis with Hurry, Fonagy appears to have regarded her as a defining exemplar of this style of interaction.

Mr S

After a degree in psychology, Fonagy pursued graduate studies in neuropsychology at University College London (UCL). He then took up a post as Lecturer in Clinical Psychology at UCL and trained as a clinical psychologist, qualifying in 1980. After qualification, Fonagy saw private clients and worked sessions as a clinical psychologist at the Royal Free Hospital. He also began training as a psychoanalyst with the British Psychoanalytic Society, graduating in 1985. He recalls that he decided to train as an analyst partly as a result of his own positive experiences in therapy with Hurry, and partly because he saw 'something within it that was a deeper explanation of how the mind functions.'⁶³ There is little in print about this period of Fonagy's life. However, a central part of psychoanalytic training with the British Psychoanalytic Society is that trainees must conduct the analysis of two patients—one for two years, one for a year—with supervision from a senior clinician. These are termed a psychoanalyst's 'membership cases', on which an oral report is given to the Society. Often the report on one of the membership cases is published. This was the case for Fonagy, offering an account of a piece of his early clinical work.

Theories in psychology, especially clinical theories, are frequently developed from some prime examples. The identification of new patient groups, as Target has noted, contributes to the development of a distinct clinical theory, reflecting the kinds of clinical examples they primarily have in mind: 'each new area of clinical interest has tended to create a new theoretical approach: narcissism begat self psychology, severe personality disorder brought forth object

⁶⁰ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 385–418, p. 393.

⁶¹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 266. Humour can offer a means of negotiating the dialectical interplay between the patient's wishes and their needs in a therapeutic relationship, allowing features of the relationship to be both real and pretend to certain degrees: Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81: 853–873: 'For the relationship to serve a function and to be tolerable, the analyst has to become what the patient wishes him to be. But at these moments, he is likely to be too terrifying for his help to be accepted. The analyst must be anything but what is projected on to him. Unless he is able to adopt an attitude analogous to that of the parent engaged in pretend play with a child, constantly juggling responses to these opposing pressures, the analysis is doomed' (pp. 868–869).

⁶² Bateman, A. and Fonagy, P. (2019). 'Introduction', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 3–20, p. 3.

⁶³ BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

relations theory, unmedicated psychotic patients spurred Kleinian thinking.’⁶⁴ It is notable, then, that the first of Fonagy’s two ‘membership cases’ with the British Psychoanalytic Society was also a patient with significant features of BPD, a new diagnosis introduced in the *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edition (DSM-III) in 1980. Yet, as well as helping direct Fonagy’s attention to the properties of BPD, the patient showed symptoms from across various diagnostic categories, suggesting the need for cross-cutting theory. Looking back on the 1980s at the early development of his thinking, for Fonagy the importance of this case cannot be overstated. It was ‘this early experience [that] led to a focus on the thinking capacities of borderline patients.’⁶⁵ Between 1989 and 2009, he offered a detailed description of the case on three different occasions. This provides an opportunity to see the way in which the case contributed to his emerging theory, and—reciprocally—the way that developments in his theory over the 20 years contributed to elaborations in his perception of the case.⁶⁶

Mr S was 27 years old when he was referred to Fonagy. He had previously been in psychotherapy for several years, but had become physically threatening to his therapist, who felt it necessary to make a referral. Mr S reported severe physical abuse as a child. His back was scarred by the early beatings. He recalled especially that he would receive physical punishment for crying. His mother had attempted suicide when he was three. Some years into the analysis, he also acknowledged that he had been sexually assaulted by his father, including on one occasion when he had gone to his father for comfort after a nightmare.⁶⁷

The primary symptoms Mr S reported at the commencement of the analysis were interpersonal problems, overwhelming anxiety, profound depression, and ‘transient psychotic-like episodes, marked by minor visual illusions.’⁶⁸ He frequently used alcohol as an attempt to regulate his wildly fluctuating emotional state. Fonagy initially attempted to offer therapeutic treatment based on psychoanalytic interpretations. For instance, he presented the interpretation that Mr S’s reluctance to discuss sexual and aggressive fantasies was due to concern about provoking a negative reaction. Fonagy recalled that ‘his response still echoes in my ears: “Don’t you understand anything? It doesn’t matter a damn what you feel. As far as I am concerned you don’t exist.”’⁶⁹ On other occasions, interpretations would

⁶⁴ Target, M. (2002). ‘Book Review’, *Symbolization: Proposing a Developmental Paradigm for a New Psychoanalytic Theory of Mind* by Anna Aragno. Madison, CT: International Universities Press, 1997, *Journal of the American Psychoanalytic Association*, 50(1): 348–352, p. 349. See also Fonagy, P. and Leuzinger-Bohleber, M. (2012). ‘Foreword’, in P. Fonagy, M. Leuzinger-Bohleber, D. Taylor, and H. Kächele (eds), *The Significance of Dreams: Bridging Clinical and Extraclinical Research in Psychoanalysis*, London: Karnac Books, pp. 3–8: ‘Some of the different psychoanalytic models of development may be rooted in these genetic differences between infants. They may all be accurate descriptions of genetically very different individuals’ (p. 151).

⁶⁵ Fonagy, P. and Bateman, A. (2009). ‘A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice’, in M. Brownescombe Heller and S. Pollet (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176, p. 159.

⁶⁶ Fonagy, P. (1989). ‘On Tolerating Mental States: Theory of Mind in Borderline Patients’. *Bulletin of the Anna Freud Centre*, 12: 91–115; Fonagy, P. (1991). ‘Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient’. *The International Journal of Psychoanalysis*, 72: 639–656; Fonagy, P. and Bateman, A. (2009). ‘A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice’, in M. Brownescombe Heller and S. Pollet (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176.

⁶⁷ Fonagy, P. (1991). ‘Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient’. *The International Journal of Psychoanalysis*, 72: 639–656, p. 649.

⁶⁸ *Ibid.* 642.

⁶⁹ Fonagy, P. (1989). ‘On Tolerating Mental States: Theory of Mind in Borderline Patients’. *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 101.

evoke no backlash, but seemed to be counterproductive and contribute to Mr S feeling suicidal.⁷⁰

Over time, Fonagy came to the conclusion that the abuse experienced by the patient had undermined his capacity to understand the thoughts and feelings of others—as well as thoughts and feelings of his own: ‘his development of adequate social understanding was impaired by trauma.’⁷¹ In fact, interpretations were experienced as assaults. The importance of impaired social understanding had been highlighted by the London Kleinian group, especially Bion, Joseph, and Britton.⁷² Fonagy agreed on the importance of this theme. But within the British Psychoanalytic Society there was terrible animosity between the Anna Freudian and Kleinian traditions still in the 1980s and early 1990s. In reflecting on the role of trauma in impairing social understanding, he distanced himself from Britton and tended to avoid citing Bion and Joseph. Somewhat caricaturing their arguments, he declared himself unpersuaded by their claim that lack of capacity for social understanding, as in Mr S’s case, should be regarded as an attack on the fantasized image of parental intercourse or on the mother.⁷³ Yet, though he criticized Britton’s interpretation, the mechanism for the impairment of Mr S’s understanding of his own and other minds was far from evident. Fonagy attempted various speculations between 1989 and 1991. These included that it represented a rejection of his mother for her lack of love, or an identification with Mr S’s ‘thoughtless’ parents, or a wish for omnipotence, or an attempt to render the analyst helpless as Mr S had been helpless as a child.⁷⁴ Even if he distanced himself from the London Kleinians regarding the origin of the problem, Fonagy found himself in agreement with Bion and Joseph that the effect of the impairment in social understanding was that Mr S was unable to make use of psychoanalytic interpretations, which lean on a patient’s capacity to recognize and think about a fundamental feature of therapy: the difference between the patient’s personal wishes and worries about their analyst, and the analyst’s here-and-now reality as an ally in reflection.⁷⁵ In fact, Fonagy believed that Mr S had only a flickering sense of other people as capable of different thoughts, perceptions, and intentions from his own, a predicament that contributed to his anxiety, anger, and bafflement in social situations such as analysis.

⁷⁰ Fonagy, P. and Bateman, A. (2009). ‘A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice’, in M. Brownescombe Heller and S. Pollet (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, p. 158.

⁷¹ Fonagy, P. (1989). ‘On Tolerating Mental States: Theory of Mind in Borderline Patients’. *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 100.

⁷² Sandler, a major influence on Fonagy, had drawn attention to the vital importance of understanding as a mental process. However, he had not addressed difficulties with social understanding specifically, as did the London Kleinians: Sandler, J. (1976). ‘Dreams, Unconscious Fantasies and Identity of Perception’. *International Review of Psycho-Analysis*, 3: 33–42.

⁷³ Bion, W. R. (1965). *Transformations: Change from Learning to Growth*, London: Tavistock; Joseph, B. ([1981] 1989). ‘Defence Mechanisms and Phantasy in the Psychoanalytical Process’, in M. Feldman and E. B. Spillius (eds), *Psychic Equilibrium and Psychic Change: Selected Papers of Betty Joseph*, London: Routledge, pp. 116–126; Joseph, B. (1983). ‘On Understanding and Not Understanding: Some Technical Issues’. *The International Journal of Psychoanalysis*, 64: 291–298; Britton, R. (1989). ‘The Missing Link: Parental Sexuality in the Oedipus Complex’, in J. Steiner (ed.), *The Oedipus Complex Today*, London: Karnac Books, pp. 83–102. For distancing from Britton, see e.g. Fonagy, P. and Target, M. (1995). ‘Understanding the Violent Patient: The Use of the Body and the Role of the Father’. *The International Journal of Psychoanalysis*, 76: 487–501, p. 497. As Bion observed, Kleinian language is somewhat metaphorical, and should not be taken in over-concrete terms. Fonagy rather makes a straw man of Britton by an overliteral interpretation of his claims.

⁷⁴ Fonagy, P. (1991). ‘Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient’. *The International Journal of Psychoanalysis*, 72: 639–656, pp. 645–646.

⁷⁵ Fonagy, P. (1989). ‘On Tolerating Mental States: Theory of Mind in Borderline Patients’. *Bulletin of the Anna Freud Centre*, 12: 91–115: ‘He clearly had difficulty in learning to differentiate between an internal transference image of his analyst on the one hand and me as a person on the other. There was no “as if” character to this transference because at times at least he had no access to a sense “real yet not real”’ (p. 103).

Additionally, Fonagy suspected that such confused perceptions about self and others contributed to the psychotic experiences of the patient.

As we have seen, Anne Hurry had already supplemented use of psychoanalytic interpretations with empathetic 'developmental help' in her treatment of adolescents. And Fonagy would later recall that, seeing the ineffectiveness of psychoanalytic interpretations of Mr S's symptoms, he thought about the clinical advice of members of the Group for the Study of Borderline Cases to avoid such interpretations when patients showed borderline symptoms.⁷⁶ Fonagy found that helping Mr S consider the source and content of his affects was a more effective intervention than psychoanalytic interpretations, able to contribute to a working collaboration and some relief of symptoms. For example, Mr S's symptoms of depression were found to centre around feelings of emptiness. Fonagy identified that these feelings seemed especially promoted by the patient's lack of vivid experience of his own or other minds. Fonagy's impression was that the most important aspect of discussions of feelings such as emptiness with the patient was not the patient's insight, or feelings of being understood, but 'the opportunity within the analytic process to entertain the way in which my mind works'. The benefit of this opportunity stemmed from the fact that 'imposing mental terms between experience and response provided him with the requisite distance to achieve control over some of his overwhelming affects'.⁷⁷ This readmitted dynamism, vividness, and a degree of choice into Mr S's sense of events and interactions, which had otherwise been dominated by certainties based on the equation of his personal impressions with reality.

Fonagy's acknowledgement of the limitations of psychoanalytic interpretation as clinical intervention fed into his approach to other early clinical cases. For instance, Mr T was a 30-year-old composer with a physical disability who entered analysis complaining of feelings of worthlessness and out-of-control aggression following alcohol use. He was cruel to himself and to others, and found it difficult to engage in cooperative activities, including conversation in clinical sessions. Fonagy attributed Mr T's difficulties in social understanding especially to his relationship with his mother, who had been cold and denigrating of him and his disability. Fonagy suspected that, despite the differences from the physical abuse experienced by Mr S, a similar outcome had occurred for Mr T in which his capacity for attending to the thoughts and feelings of himself and others had been inhibited. Specifically, in the case of Mr T, his mother's treatment had led him to use his disabled body 'as the theatre of his mental experience', rather than having thoughts and feelings serve as this theatre.⁷⁸ Usually, the body is

⁷⁶ Rosenfeld, S. K. and Sprince, M. P. (1965). 'Some Thoughts on the Technical Handling of Borderline Children.' *The Psychoanalytic Study of the Child*, 20(1): 495–517. Fonagy, P. and Bateman, A. (2009). 'A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice', in M. Brownescombe Heller and S. Pollet (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176: 'I began to dimly remember Rosenfeld's caution about interpreting unconscious conflict with patients with borderline personality disorder. My unease increased when he went on to remember a further segment of the dream when the image of castrating his father was quickly followed by seeing himself in the dream holding his father's severed penis up as if it were a torch. He was reminded of the Statue of Liberty and conveyed to me that illuminating the darkness represented by his unconscious mind freed him to be a person who could now genuinely engage with others. Needless to say, neither my interpretations nor his dream the following day signalled a dramatic improvement. On the contrary, he appeared to deteriorate to become even more suicidal' (p. 158). In a late lecture, Winnicott had also reflected that 'psychoanalysis is not just a matter of interpreting the repressed unconscious; it is rather the provision of a professional setting for trust, in which such work may take place.' Winnicott, D. W. ([1970] 1986). 'Cure', in *Home Is Where We Start From: Essays by a Psychoanalyst*, New York: Norton, pp. 112–122, p. 115. Though he does not anywhere cite this passage, Winnicott had a tremendous general influence on Fonagy's early thinking.

⁷⁷ Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients.' *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 106.

⁷⁸ Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration.' *Bulletin of the Anna Freud Centre*, 17: 33–55, p. 51.

part of the archaic accompaniment to language, conveying tone but not asked to stand alone for meaning. However, 'if objects cannot be properly represented as thinking and feeling, they may to some extent be controlled, distanced or brought into proximity through bodily experiences'.⁷⁹ As for Mr S, the difficulties with social understanding necessitated changes in clinical technique:

As he clearly found it hard to stick to the normal constraints of the analytic encounter, I decided also to permit myself occasional deviations from 'classical' technique and act simply as any human being might be expected to react to such provocations. For example, when he peered down on me, standing above my chair, I simply stood up from my chair and said 'You are trying to make me feel small, but you know I am actually taller than you are' . . . My interventions served to restore contact with him at a human level, and on most occasions he would regain his composure, and some kind of analytic dialogue would resume.⁸⁰

By the early 1990s, Fonagy's experiences had oriented him towards a concern with perspective-taking and its clinical implications. His father's interests and his own self-reflection had highlighted that social understanding in adulthood is grounded upon the base of childhood experiences, and that this contributed to the dynamic interplay of meanings in adult speech and social understanding. His father's diary had documented interest in the child's potential to confuse internal and external perception, confuse real and pretend, and lose the difference between concrete occurrences and their different potential meanings. However, Fonagy had contested his father's sense that these represented a child's confusion, instead proposing that children may process the world in a qualitatively different way, and that distinctions familiar to adults are developmental achievements. Such achievements may be interrupted briefly or blocked off in a more sustained way depending on the care a child has received, and especially depending on whether or how parents have offered acknowledgement of the child's experience.

These concerns shaped his early clinical work and contributed to his interest in his patients' social understanding of thoughts and feelings. From technical innovations associated with the Group for the Study of Borderline Cases and from his own analysis as an adolescent with Anne Hurry, Fonagy was primed to accept the limitations and problems of traditional psychoanalytic interpretation as a form of clinical intervention. Instead, he had begun to experiment with clinical interventions focused around helping patients gain perspective on the sources and scope of their feelings and thoughts. These reflections would be further developed by his work as a clinician, researcher, and later manager at the Anna Freud Centre.

⁷⁹ Fonagy, P. and Target, M. (1995). 'Understanding the Violent Patient: The Use of the Body and the Role of the Father'. *The International Journal of Psychoanalysis*, 76: 487–501, p. 495.

⁸⁰ Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration'. *Bulletin of the Anna Freud Centre*, 17: 33–55, pp. 40–41.

2

Work at the Anna Freud Centre

Introduction

Chapter 1 presented three biographical snapshots of some of Fonagy's formative experiences: as a child in the Fónagy household, as an adolescent in therapy with Hurry, and as a young analyst in the 1980s. This chapter will continue chronologically into the 1990s, but with a greater focus on institutional rather than biographical context. It will explore Fonagy's work as research director for the Anna Freud Centre in the 1990s, and the research conducted by Target and Fonagy examining the long-term outcomes of former patients at the Centre.

In considering the institutional context within which Fonagy and colleagues have developed their ideas, the focus on the Anna Freud Centre is in part an effect of the available textual record. Fonagy has stated in interview that, after the suicidal feelings and depression of his adolescence, 'these two institutions saved me: both the Anna Freud Centre and University College London (UCL).'¹ However, there is almost nothing in the public domain documenting Fonagy's activity within UCL. This is despite the fact that he has spent his whole career at UCL: as undergraduate, graduate, lecturer, head of department for Clinical, Education, and Health Psychology since 2008, and as head of division for Psychology and Language Sciences since 2017.² That said, Fonagy has expressed pride in the UCL tradition of 'open-mindedness', as the first English university to accept atheists and religious non-conformers, and then later the first to accept women.³ He has likewise signalled his sympathy for the utilitarian values of the founder of UCL, Jeremy Bentham, for whom the purpose of knowledge was its value in practical use.⁴

After describing Fonagy's work at the Anna Freud Centre in the 1990s, this chapter will attempt to characterize some features of Fonagy's leadership of the Centre since 2003. The Centre has seen an incredible transformation in this time. We will seek to situate this transformation in relation to the challenges and opportunities of the wider social context. The chapter will close with an attempt to briefly situate a number of the major collaborations from different eras of Fonagy's work, introducing the *dramatis personae* for the rest of the book.

Research director at the Anna Freud Centre

After graduating as a psychoanalyst accredited to practise with adults in 1985, Fonagy pursued a private practice while also working as a lecturer in psychology at UCL. In 1989, he

¹ Fonagy, P. (2018). 'Peter Fonagy: Combating a Mental Health Crisis'. Accessed at: <https://www.goldmansachs.com/insights/talks-at-gs/peter-fonagy.html>.

² Fonagy, P. (2017). 'A Word from our New Head of Division Professor Peter Fonagy to all PALS Alumni'. Accessed at: <https://www.ucl.ac.uk/pals/sites/pals/files/peter-fonagy-letter-to-alumni.pdf>.

³ Fonagy, P. (2014). 'Why UCL?'. Accessed at: <https://www.youtube.com/watch?v=XbvKbHO57ZI>.

⁴ Ibid.

began training as a child and adolescent psychoanalyst at the Anna Freud Centre. His clinical supervisors were Rose Edgumbe, Anne-Marie Sandler, and Marion Burgner. As we saw in Chapter 1, he had already come to identify limitations to interpretation as the primary form of clinical intervention, which characterized the 'classical' psychoanalytic technique. His critical appraisal of the inherited psychoanalytic tradition was not always encouraged:

You know, I nearly didn't make it through my training. Halfway through I was pulled up by my progress advisor, who asked me if I was certain that analysis was the career for me. I asked why, and she said that it was because I had been critical in seminars. It's not that my career is to criticize, I told her, I just wanted to test the ideas, and I tested them.⁵

Nonetheless, Fonagy was supported by his clinical supervisors to utilize the technique of 'developmental help' (see Chapter 1) in his work with child patients. Through the early 1990s, he reported cases in which it appeared to be this aspect of clinical technique that proved effective at reducing his child patients' symptoms and helping them to understand themselves and others.⁶ For instance, Fonagy reported the case of William, a 7 year old who could not relate to or play with his peers. His mother wished that he had never been born and made him feel rejected. This was disguised by a barrage of talk to and about him, but which showed little recognition of his intentions, thoughts, or feelings. Neither she nor William's father had played with him. Fonagy reported:

After three years of intensive 'developmental help', William was far more able to think about himself and others in terms of motives and feelings, and to explore the painful area of how different his thinking was from other people's and how this had cut him off from them. Throughout the analysis, play, focused on uncovering the psychic reality behind the physical, and then on the different perspectives of different people, formed the backbone of all the therapeutic work.⁷

As well as pursuing training in child and adolescent analysis, from 1989, Fonagy was appointed research director at the Centre. One of his first projects was to put together a technique manual for child psychoanalysis at the Centre. This was undertaken in conjunction with Mary Target and Rose Edgumbe, and as the doctoral project for Jill Miller. Initially, work on this manual was oriented by the principles of classical psychoanalytic technique for work with children and adolescents, with a focus on reticence and psychoanalytic interpretation of both patients' symptoms and the dynamics in the patient–therapist relationship. However, there was a growing consensus that developmental help was of special importance for the therapy provided by the Anna Freud Centre. So Fonagy and colleagues started again,

⁵ Fonagy, P. (2007). 'Interview', in L. E. Rubinstein (ed.), *Talking about Supervision: 10 Questions, 10 Analysts = 100 Answers*, London: International Psychoanalytic Association, pp. 39–49, pp. 45–46.

⁶ The 'developmental help' approach at the Anna Freud Centre found support also from Paulina Kernberg's object relations approach, though it would not appear that Kernberg had significant direct influence on Fonagy. Kernberg, P. F. and Chazan, S. E. (1991). *Children with Conduct Disorders: A Psychotherapy Manual*, New York: Basic Books. Another, apparently independent voice from the period calling for a shift from interpretation to supportive counselling was Killingmo, B. (1989). 'Conflict and Deficit: Implications for Technique'. *International Journal of Psychoanalysis*, 70: 65–79. The Sandler recall that, at least in the UK context, a focus on relational support among Anna Freudians was influenced by work in the independent tradition (e.g. Winnicott). Sandler, J. and Sandler, A. M. (1994). 'The Past Unconscious and the Present Unconscious: A Contribution to a Technical Frame of Reference'. *The Psychoanalytic Study of the Child*, 49(1): 278–292, p. 280.

⁷ Fonagy, P. (1995). 'Playing with Reality: The Development of Psychic Reality and its Malfunction in Borderline Personalities'. *The International Journal of Psychoanalysis*, 76(1): 39–44, pp. 42–43.

this time with developmental help as the focus. Again, however, this attempt failed: there was too much dissensus for a manual to be developed that met with enough approval to publicly represent the Centre.⁸ Part of the problem was that ‘the same words were being used in significantly different ways by different clinicians, who had always assumed that they were talking and writing about the same phenomena.’⁹ Despite this, Fonagy was struck that amid the dissensus that blocked the publication of the manual ‘we found that we had use for very few terms. As for all the subtlety—you just don’t need it.’¹⁰

In the abandoned manual, Fonagy and colleagues attempted to characterize the core elements of developmental help. On the one hand, developmental help offers support to the child or young person’s ‘dare to change’, for instance—by drawing the patient’s attention to limitations of their current strategies and the possibilities held out by alternatives. In doing so, the therapist is encouraged to act with due acknowledgement that this may be quite challenging for patients whose symptoms reflect adaptations to an ‘environment perceived as dangerous or destructive in which improvement in his functioning would not be welcome.’¹¹ On the other hand, the techniques of development offer direct assistance to the patient to support successful psychological functioning, and thereby ‘increase the child’s sense of being competent’. This may include ‘support in learning to control his feelings and behaviour, suggestions or demonstrations of how to manage difficulties he encounters, and so on’, and should be accompanied by a good deal of ‘encouragement and praise.’¹² The therapist acts as an ‘auxiliary’ source of help for the patient’s capacities for affect regulation and behavioural control, encouraging the young person as they learn to achieve these capabilities independently.

The ideas circulating at the Anna Freud Centre during Fonagy’s training had particular importance for his later thinking in four regards. A first was the growing interest at the Centre in self-understanding and the experience of being understood to clinical work in general, but especially to work with children and adolescents and in patients presenting with borderline states. A second was the perspective that the most crucial aspect of ‘borderline’ phenomena was difficulties in relating to others, and that this may stem from trauma in the child–caregiver relationship. A third was the idea of developmental help, seeking to help young people find new forms of adaptation to their environment that would permit greater affect regulation and self-esteem. And a fourth important lesson, related to the work of Fonagy and colleagues on the manual, was that clinicians might have significant overlap in the core elements that contribute to successful therapy.

In these and other ways, the emerging stance of Fonagy and his immediate collaborators in the 1990s was indebted to the Anna Freudian tradition. However, Fonagy and colleagues appear to have experienced this tradition as ultimately no longer adequate for its contemporary tasks. Some aspects, alive and bright, were pulled out and preserved; others were

⁸ Edgumbe, R. (1995). ‘The History of Anna Freud’s Thinking on Developmental Disturbances’, *Bulletin of the Anna Freud Centre*, 18(1): 21–34: ‘As recently as five or six years ago, when Peter Fonagy, George Moran, Hansi Kennedy and I started trying to put together a Technique Manual, we began by trying to orientate it around classical technique with neurotic children. But we and the therapists whom we asked to comment on our formulations found this unsatisfactory and we were obliged start again, this time trying to formulate developmental help. But we failed to sort it out satisfactorily’ (p. 22).

⁹ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 289.

¹⁰ Fonagy, P. (1999). ‘Interview with Peter Fonagy’, in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 88.

¹¹ Fonagy, P., Edgumbe, R., Target, M., and Miller, J. (1999). *Contemporary Psychodynamic Child Therapy: Theory and Technique*, London: The Anna Freud Centre and University College London, p. 111, unpublished manuscript, Mary Target’s personal archive.

¹² *Ibid.*

patched with grafts from elsewhere. Anna Freud was centrally interested in the role of child-caregiver relationships and in education for providing scaffolding for an individual's skills at self-regulation, and as the locus for preventative interventions.¹³ Both concerns were retained by Fonagy and colleagues. Anna Freud's interest in both positive and negative factors in development, and a child's adaptation to their circumstances, would be pivotal to the work of Fonagy and his collaborators. Anna Freud also tended to eschew psychiatric diagnoses as superficial, in favour of a quasi-interval scale of total mental health needs. This can be regarded as a relevant backdrop to Fonagy and colleagues' critique of diagnosis-focused mental health practice, and their conceptualization of mental illness as a latent dimension.¹⁴

Yet the position of Fonagy and colleagues also represented an important and valuable advance, leading to a substantially more plausible theory. Several rather weak and speculative aspects of the Anna Freudian tradition were rejected, including Anna Freud's use of drive theory and her appeal to the id, ego, and superego as if they were distinct agencies.¹⁵ Freud's analysis of development in terms of oral, anal, and genital stages was also, appropriately, thrown out. Perhaps the most influential shift, however, was in the conceptualization of the psychological processes that generate positive and negative mental health. Anna Freud had emphatically argued for the therapeutic priority of psychological processes that serve as defence mechanisms against anxiety, unpleasant realities, and the conflict of incompatible drives. For Anna Freud, defences—for instance, fantasies of omnipotence, or workaholism as a sublimation—may contribute to pathology, or to positive adaptation, depending on the subtlety and sophistication of the defence, and its responsiveness to the demands of the environment.¹⁶ By contrast, Fonagy and colleagues advocated the therapeutic priority of psychological processes that obstruct the capacity to conceive of or reconsider thoughts and feelings. They have also conceptualized resilience as the capacity to learn from experience. So, for instance, the concept of repression, foundational for Freud, was abandoned by Fonagy and colleagues in favour of a revised account of distortions in cognitive and affective processing.¹⁷ Even when the accounts could in principle have been integrated—for instance, in considering what fantasies are helpful in facilitating the capacity to conceive of or reconsider

¹³ Freud, A. (1954). 'Psychoanalysis and Education'. *The Psychoanalytic Study of the Child*, 9(1): 9–15; Freud, A. (1965). *Normality and Pathology in Childhood*, New York: International University Press.

¹⁴ Anna Freud also at times expressed scepticism of the role of single traumas in mental health symptoms, and argued—as Fonagy and colleagues would later—that accounts of trauma in therapeutic contexts may be important but in large part as narrative constructions through which a patient works to understand and recalibrate their experience of themselves. However, Freud was not consistent on this point. Freud, A. (1958). 'Child Observation and Prediction of Development: A Memorial Lecture in Honor of Ernst Kris'. *The Psychoanalytic Study of the Child*, 13(1): 92–116. Another intellectual ancestor of Fonagy who may have provided a relevant backdrop was Winnicott, for whom mental ill health was arrayed quantitatively, with no sharp qualitative lines between diagnoses—even between health and full-blown schizophrenia. Winnicott, D. (1971). *Playing and Reality*, London: Routledge, p. 66.

¹⁵ Midgley, N. (2012). *Reading Anna Freud*, London: Routledge.

¹⁶ The 10 defences privileged by Anna Freud were repression, regression, reaction-formation, isolation, undoing, projection, introjection, turning against the self, reversal into the opposite, and sublimation. Freud, A. (1946). *The Ego and the Mechanisms of Defense*, New York: International Universities Press. Some of these processes may be facilitated or actually achieved through forms of non-mentalizing (see Chapter 5). Regression, reaction-formation, and reversal into the opposite may be prompted by pretend mode. Undoing, projection, and introjection may be prompted by psychic equivalence, in which what is *felt* is experienced as real (e.g. that one can really undo an action through some symbolic reparation). Isolation may be achieved by teleological mode. However, the defence mechanisms and the forms of non-mentalizing operate on rather different levels of analysis, given that they are descriptions of the obstruction of mental processes that are conceived very differently. Exemplary in this regard is the causal centrality of anxiety for Freud, and its peripheral and epiphenomenal role for Fonagy and colleagues, in their conceptualization of the mind and of mental illness.

¹⁷ Fonagy, P. (2000). 'Response'. *The International Journal of Psychoanalysis*, 81(2): 354–356. See also Pedersen, S. H. (2013). 'Fonagy and Freud. Psychological versus Psychic Reality'. *The Scandinavian Psychoanalytic Review*, 36(1): 18–26.

thoughts and feelings, or in conceptualizing and working clinically with anxiety—such questions have been left essentially untouched, because the concepts for asking them have been set aside.¹⁸

Fonagy and colleagues stitched together work at the Anna Freud Centre with ideas from other traditions. This included the treachery of claiming commonalities and opportunities for mutual learning between the Anna Freudians and their traditional arch-enemies, the Kleinian school of psychoanalysis. For instance, Fonagy felt that the Kleinians had made major advances in studying the patient's awareness of the clinician's mind, and the patient's experience of reality in patient–analyst interaction.¹⁹ He adopted a version of the Kleinian concept of 'projective identification' in the idea of the externalization of the alien self (see Chapter 6).

Fonagy also asserted opportunities for mutual learning between psychoanalysis and non-psychoanalytic approaches to supportive and therapeutic work with children and adults. Fonagy went on record praising Cognitive Behavioural Therapy (CBT), a modality that was increasingly coming to dominate publicly funded mental health care, to an extent at the expense of psychoanalysis. Though he regarded CBT as offering less depth and richness than psychoanalysis, he felt that it had introduced new techniques and values that might be of general value to therapeutic work: 'I think cognitive behavioural therapy is a fairly rigorous business and I have a lot of time for it. The therapists often know what they're talking about and, if done well, CBT can be very good.'²⁰ Fonagy knew he was courting controversy in making such assertions. In interview, he would describe his self-perception as follows:

I'm Peter Fonagy. I consider myself to be a Freudian. But in fact, it is the Freudians that I have somewhat greatest difficulty with. And many Freudians now consider me a Kleinian. And those who do not consider me a Kleinian consider me independent. And those who do not consider me independent consider me a renegade.²¹

My theoretical stance is really a distortion of Anna Freud's thinking, although I know that, if she was alive, she wouldn't think I was following her ideas. But I certainly feel that it's really Anna Freud's ideas that I'm trying to develop.²²

¹⁸ Another point of potential integration could have been around Part II of Anna Freud's *Ego and the Mechanisms of Defence*, which is centrally concerned with strategies to avoid social understanding by blocking occasions for learning from experience. So, for instance, a new defence mechanism is proposed by Anna Freud—'restriction of the ego'—in which an individual structures their social environment to pre-emptively avoid unpleasure. However, Freud's examples are largely of strategies to pre-emptively avoid certain thoughts and feelings about themselves or others. This suggests a role for the affordances of the environment in facilitating or hindering mentalizing that Fonagy and colleagues have themselves been concerned with, especially in recent years (see Chapter 9). However, the difference in metapsychology, and particularly Freud's use of drive theory, has made her account of defences somewhat difficult for Fonagy and colleagues to play with and use, or to subsequently revisit. Freud, A. (1946) *The Ego and the Mechanisms of Defence*, New York: International Universities Press.

¹⁹ E.g. Joseph, B. (1985). 'Transference: The Total Situation'. *The International Journal of Psychoanalysis*, 66: 447–454; Feldman, M. (1993). 'Aspects of Reality, and the Focus of Interpretation'. *Psychoanalytic Inquiry*, 13: 274–295; Steiner, J. (1993). *Psychic Retreats*, London: Routledge. Fonagy and colleagues express enthusiasm for these works in Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 448.

²⁰ Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 93.

²¹ Palmer, S. (2015). 'Controversial Discussions for the XXIst Century: An Interview with: Jan Abram, Dana Birkstead-Breen, Catalina Bronstein, Peter Fonagy, Bob Hinshelwood, Isabel Hernandez-Halton, Rosine Perelberg, Ken Robinson, Anne-Marie Sandler, Allan Schore, Mark Solms, Riccardo Steiner, and David Tuckett'. PEP Video Grants, 1(1): 2.

²² Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 96.

One quality that set Fonagy apart from most other Anna Freudians, and contributed to their perception of him as a potential threat, was his concern with the scientific evaluation of psychoanalytic practice. This external standard was one in which most psychoanalytic practitioners had little skill or training. It also challenged established regimes of power and authority, especially the right to judge what counted as reality. Scientific evaluation was perceived by many of Fonagy's colleagues as a profanation, with the result that 'analytic therapists in particular have been defensive and suspicious in the face of the evidence culture as though the very idea of objective scrutiny represented a hostile intrusion into a quasi sacred private world.'²³ Sigmund Freud had, at times, claimed that any element of psychoanalysis could be jettisoned if subsequent scientific developments suggested the need to update theory or practice.²⁴ However, Freud had little interest in experimental studies as a means of evaluating and improving psychoanalytic practice.²⁵

With a few exceptions such as John Bowlby,²⁶ by the time Fonagy was training as an analyst, psychoanalysts had little appetite for either experimental research or the incorporation of new developments from other scientific disciplines. Instead, Fonagy saw around him that developments in psychoanalytic theory tended to stem from reflection on individual cases or a handful of cases, with ideas incorporated into the canon on the basis of the appeal to other clinicians, and without established systems to prune or disconfirm theory:

The abundance of clinicians claiming, on the basis of case reports, that their theory and technique are indispensable, is the gravest indictment of the logic of case-study methodology. It leaves open the possibility that an unspecifiable, but possibly very substantial, portion of coherent psychoanalytic generalizations, which meet the hermeneutic criteria of consistency and meaningfulness, are incorrect.²⁷

Allied with Joseph Sandler's concern regarding the conflation of description and explanation, and with work pursued by Ivan Fónagy on the nature of metaphor in science,²⁸ Fonagy argued in the 1980s that 'all the metaphoric language of psychoanalysis is a sign of inadequate information concerning underlying psychological processes. It indicates gaps in our knowledge which will only be filled by further experimental research.'²⁹ No denigration of

²³ Goldbeck-Wood, S. and Fonagy, P. (2004). 'The Future of Psychotherapy in the NHS'. *British Medical Journal*, 329: 245–246, p. 245. See also Chiesa, M. and Fonagy, P. (2010). 'Scientific Research, the Therapeutic Community and Psychodynamic Psychotherapy'. *Clinical Neuropsychiatry*, 7(6): 173–181.

²⁴ Freud, S. ([1940] 2001). *An Outline of Psycho-Analysis (standard edn)*, Volume 23, London: Vintage, pp. 144–207, p. 159.

²⁵ Shulman, D. G. (1990). 'Psychoanalysis and the Quantitative Research Tradition'. *Psychoanalytic Review*, 77(2): 245–261.

²⁶ Bowlby, J. (1981). 'Psychoanalysis as a Natural Science'. *The International Review of Psychoanalysis*, 8: 243–256; Wallerstein, R. (1986). 'Psychoanalysis as a Science: A Response to the New Challenges'. *Psychoanalytic Quarterly*, 55(3): 414–451.

²⁷ Fonagy, P. and Tallandini-Shallice, M. (1993). 'On Some Problems of Psychoanalytic Research in Practice'. *Bulletin of the Anna Freud Centre*, 16: 5–22, p. 6. See also Fonagy, P. (2000). 'On the Relationship of Experimental Psychology and Psychoanalysis: Commentary by Peter Fonagy (London)'. *Neuropsychanalysis*, 2(2): 222–232: 'Clinical data clearly offer a fertile ground for theory building, but not for distinguishing good theories from bad or better ones' (p. 228).

²⁸ Fónagy, I. (1989). 'The Metaphor: A Research Instrument', in D. Meutsch and R. Viehoff (eds), *Comprehension of Literary Discourse*, Berlin and New York: W. De Gruyter, pp. 111–130.

²⁹ Fonagy, P. (1982). 'The Integration of Psychoanalysis and Experimental Science: A Review'. *International Review of Psycho-Analysis*, 9: 125–145; Sandler, J. (1983). 'Reflections on Some Relations between Psychoanalytic Concepts and Psychoanalytic Practice'. *The International Journal of Psychoanalysis*, 64: 35–45. This point would later be repeated and elaborated in Luyten, P. (2015). 'Unholy Questions about Five Central Tenets of Psychoanalysis that Need to be Empirically Verified'. *Psychoanalytic Inquiry*, 35(1): 5–23. For a sustained discussion of the value of metaphor, see Civitarese, G. and Ferro, A. (2013). 'The Meaning and Use of Metaphor in Analytic Field Theory'.

metaphor was intended; Fonagy himself clearly appreciated the value of a good metaphor in developing theory and in scientific exposition. Rather, his point was that metaphor should not be mistaken for causal-deductive claims about mental processes.

Fonagy's distinction between evocative metaphor and causal-deductive psychology was over-sharp; lack of attention to what is evoked by the language of psychological theory would contribute to confusion regarding the meaning of his own theoretical terms like 'disorganization' and 'self' (see Chapters 3 and 6).³⁰ However, in his discussion of metaphor, Fonagy is certainly persuasive in his critique of the language of psychoanalytic theory: this language may give a luminous *feeling* of comprehension, but Fonagy felt that its tendency towards encompassing metaphors masks imprecision, especially in conceptualizing and predicting causal processes. The audience nod along, and lose the sense that the ideas are provisional estimates.³¹ This is because a shortcut has been enacted in the process of understanding, because individuals take away quite different meanings from the theory based on their existing presuppositions, while believing they hold a common notion. With psychoanalytic theory both superabundant and ultimately unreliable in the articulation of causal processes, Fonagy saw clinicians operating in practice on the basis of implicit assumptions and low-grade generalizations about what was felt to work with patients, assumptions that were not brought for public discussion or elaboration into testable hypotheses. The result was 'a petrification of clinical practice,' a weak articulation of the underlying psychological processes acted upon in clinical interventions, and resistance to even thinking about the reform of how analysts are trained.³²

Supported by a collaboration with George Moran, Director of the Anna Freud Centre until his early death in 1992, Fonagy sought to pursue and nurture research at the intersection of psychoanalysis and scientific practice.³³ After completing his training as a child psychoanalyst in 1995, he co-convened yearly research training seminars in London, under the banner of the International Psychoanalytical Association, to support the development of empirical research projects by psychoanalysts and trainees, as a step towards a culture favourable to psychoanalytic research.³⁴ As chairman of the International Psychoanalytic

Psychoanalytic Inquiry, 33(3): 190–209; Havsteen-Franklin, D. (2019). 'Creative Arts Therapies', in Anthony Bateman and Peter Fonagy (eds). *Handbook of Mentalising in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 181–195. See also Holmes, J. and Slade, A. (2017). *Attachment in Therapeutic Practice*, London: Sage: 'Metaphor is inherently mentalising' (p. 70).

³⁰ For criticism of the contrast between metaphor and propositional discourse on the grounds that the latter, too, is subject to implicatures, see Sperber, D. and Wilson, D. (2008). 'A Deflationary Account of Metaphor', in Raymond W. Gibbs, Jr., (ed.), *The Cambridge Handbook of Metaphor and Thought*, New York: Cambridge University Press, pp. 84–105; Sperber, D. and Wilson, D. (2015). 'Beyond Speaker's Meaning'. *Croatian Journal of Philosophy*, 15(44): 117–149.

³¹ Bion had earlier raised an aligned concern, that psychoanalytic theory actually proves a systematic obstacle to learning from experience to the extent that the epistemic status of its language is misunderstood. However, whereas Fonagy contrasts metaphor and causal-deductive claims, Bion addresses a wider variety of 'rows' of kinds of claim, differentiating kinds of proto-thought and also kinds of abstract thought, again which he urges should not be mistaken for one another.

³² Fonagy, P. (2003). 'Some Complexities in the Relationship of Psychoanalytic Theory to Technique'. *Psychoanalytic Quarterly*, 72(1): 13–47, p. 38. See also Fonagy, P. (2009). 'When Analysts Need to Retire: The Taboo of Ageing in Psychoanalysis', in B. Willock, R. C. Curtis, and L. C. Bohm (eds), *Taboo or not Taboo?* London: Karnac Books, pp. 209–227, p. 220; Fonagy, P. (2012). 'On Caution and Courage in Psychoanalytic Epistemology'. *Philosophy, Psychiatry, & Psychology*, 19(3): 213–215.

³³ Moran, G. S. and Fonagy, P. (1987). 'Psychoanalysis and Diabetic Control: A Single-Case Study'. *Psychology and Psychotherapy: Theory, Research and Practice*, 60(4): 357–372.

³⁴ Emde, R. N. and Fonagy, P. (1997). 'An Emerging Culture for Psychoanalytic Research?' *The International Journal of Psychoanalysis*, 78(4): 643–651; Wallerstein, R. S. and Fonagy, P. (1999). 'Psychoanalytic Research and the IPA: History, Present Status and Future Potential'. *The International Journal of Psychoanalysis*, 80: 91–109. See

Association Standing Committee for Research from 1993 to 2005, Fonagy was also involved in the funding of psychoanalytic research.³⁵

It should not be thought that Fonagy's position was that of the dour empiricist: for example, he would act as an advocate for speculative psychoanalytic theory in academic contexts, arguing for recognition of its potential contribution.³⁶ He valued the inherited tradition of hermeneutic thinking in psychoanalysis as having generated many unique and valuable proposals about the human mind and about the treatment of mental suffering. He was certainly not above offering untestable speculations at times. For instance, Fonagy and Target speculated, apparently seriously, that psychoanalysts' aversion to new beliefs and knowledge from outside psychoanalysis was an 'unconscious expression of infantile patterns', which made them afraid to explore freely.³⁷ As we will see in Chapter 5 on the modes of non-mentalizing, Fonagy and colleagues have not hesitated to give clinical guidance over decades without getting round to operationalizing relevant constructs for testing. Fonagy's concern has not been with speculation per se, but with speculation intrinsically cut off from empirical testing. He felt, following Karl Popper and others, that a distinction should be drawn between the work of proposing hypotheses (the 'context of discovery') and the work of testing them enough to be able to be sure of the evidence and its replicability (the 'context of justification').³⁸ Some analysts might wish to focus their energies on proposing hypotheses and some on testing them, but psychoanalysis as a discipline needed both activities for the sake of its internal health and development, as well as for its external standing.

In the second half of the 1990s and early 2000s, Fonagy wondered at times whether it was already too late for psychoanalysis to change fast enough to counteract its reputation as a scientific backwater, and to gain the credibility to receive research funding and opportunities

also Fonagy, P. (2018). 'An Appreciation of Dr Wallerstein's Contributions to Psychoanalytic Research', in Wilma Bucci (ed.), *From Impression to Inquiry*, London: Routledge, pp. 18–21.

³⁵ André Green would criticize Fonagy for using the financial resources of the International Psychoanalytical Association to incentivize trainee psychoanalysts to pursue research. In his view, Fonagy and colleagues were setting up a new 'objectivist' school of psychoanalysis. Green, A. (2000). 'Science und Science-fiction in der Sauglingsforschung'. *Zeitschrift für psychoanalytische Theorie und Praxis*, 15(4): 438–466. Fonagy and Hepworth would retort that scientific measurement is not a 'school', but precisely a means of constraining the organization of psychoanalytic theory to avoid a proliferation of incommensurable schools. Fonagy, P. and Target, M. (1996). 'Should We Allow Psychotherapy Research to Determine Clinical Practice?' Comments on Sol J. Garfield: 'Some Problems Associated with "Validated" Forms of Psychotherapy'. *Clinical Psychology: Science and Practice*, 3: 245–250, p. 248.

³⁶ Polatinsky, S. (2005). *Psychoanalysis and the Non-Conceptual*. Unpublished doctoral thesis, University College London. Accessed at: <http://discovery.ucl.ac.uk/1445792/1/U593116.pdf>: 'My sincere thanks to Peter Fonagy for his distinctive contribution of much needed humour and irreverence, and most importantly for his unstinting interest in this research project particularly in light of (in the face of) the implacable empiricists and logical positivists holding court (and fort!) at UCL.'

³⁷ Fonagy, P. and Target, M. (2007). 'The Rooting of the Mind in the Body: New Links between Attachment Theory and Psychoanalytic Thought'. *Journal of the American Psychoanalytic Association*, 55(2): 411–456, p. 446. This position was later updated with the more plausible proposal that the psychoanalytic community had adopted a stance of epistemic vigilance, making them quite capable of generating new ideas but within a context in which they were unable to learn from outside perspectives. Allison, E. and Fonagy, P. (2016). 'When is Truth Relevant?' *Psychoanalytic Quarterly*, 85(2): 275–303.

³⁸ Emde, R. N. and Fonagy, P. (1997). 'An Emerging Culture for Psychoanalytic Research?' *The International Journal of Psychoanalysis*, 78(4): 643–651: 'Exploratory methods, aimed at discovery, can be distinguished from hypothesis-testing methods, in the context of confirmation' (p. 647). See also Fonagy, P. (2012). 'On Caution and Courage in Psychoanalytic Epistemology'. *Philosophy, Psychiatry, & Psychology*, 19(3): 213–215: 'Nietzsche talks of unpretentious truths that have been discovered by means of rigorous method, and opposes them to the metaphysics that blinds us and make us happy. He is essentially distinguishing boring Millian fact from evocative narrative. Holding on to these unpretentious truths demonstrates courage of a different sort from that shown by psychoanalytic investigations of the unconscious. It is a turning away from what is appealing toward what is true' (p. 215).

for collaboration with significant figures in other disciplines.³⁹ When he and Mark Solms attempted to establish a new MSc in neuropsychanalysis at UCL, the application was met with rejection on the basis that psychoanalysis had insufficient collaborations with neuroscience to warrant a postgraduate degree at the intersection.⁴⁰ Nonetheless, in general, Fonagy retained optimism and, in papers from 1997, urged major and immediate changes to psychoanalytic culture:

If we fail to meet the challenge confronting us, if child analysts offer their treatment for an overly extended range of disorders, while making little attempt to demonstrate its effects in a form accessible to others, the likely consequence is that child psychoanalysis will be discredited and disappear. This will be a great loss to psychoanalysts, but a still greater one to those children who specifically need this form of help but who will no longer have access to it.⁴¹

Fonagy, Bleiberg, and Target advised pursuit of cross-disciplinary collaborations and integrations of psychoanalysis with new developments in science. And, to achieve this, they called for six changes to psychoanalytic practice to generate the basis for scientific credibility.⁴²

First, Fonagy argued that psychoanalysis needed better characterization of its interventions, to the point that these could be manualized. This would clarify the relationship between intervention, process, and outcome.⁴³ A second was that better specification was needed regarding what interventions were appropriate for what symptoms ('what works for whom?'). It should not be assumed that the same techniques, or even psychoanalytic approaches in general, would be equally effective for conditions as diverse as anxiety disorders, drug addiction, and postnatal depression. Third, clinical practices needed to develop independent and quantitative evidence of their efficacy.⁴⁴ It might be a complex task to quantify the benefit a patient receives from psychoanalysis, given that this is generally assumed to be much more than the reduction in their presenting symptoms. Fonagy felt that this complexity was not an intrinsic obstacle to evaluation, though it required candid discussions about what benefits might be seen and how they might be authentically measured.

³⁹ E.g. Fonagy, P. and Target, M. (2002). 'The History and Current Status of Outcome Research at the Anna Freud Centre'. *The Psychoanalytic Study of the Child*, 57(1): 27–60, p. 56. Fonagy would later reflect on the predicament of art therapy, which he is confident is clinically effective, but which has not generated an adequate research base to now readily compete for a share of research funding. Fonagy, P. (2012). 'Art Therapy and Personality Disorder'. *International Journal of Art Therapy*, 17(3): 90. Fonagy described this as an especially sad predicament for him since 'I believe that with its many subspecialties, art therapy has the key, or perhaps a key, to our understanding of the mechanisms underpinning change in all kinds of psychological treatments ... art therapy is closest to what we now understand to be the embodied roots of human consciousness and cognition' (p. 90). See also Havsteen-Franklin, D. (2019). 'Creative Arts Therapies', in Anthony Bateman and Peter Fonagy (eds) *Handbook of Mentalising in Mental Health Practice* (pp.181–195), Washington, DC: American Psychiatric Association.

⁴⁰ Fonagy, P. (2003). 'Genetics, Developmental Psychopathology, and Psychoanalytic Theory: The Case for Ending our (not so) Splendid Isolation'. *Psychoanalytic Inquiry*, 23(2): 218–247, p. 242.

⁴¹ Fonagy, P. and Target, M. (1997). 'The Problem of Outcome in Child Psychoanalysis: Contributions from the Anna Freud Centre'. *Psychoanalytic Inquiry*, 17(S1): 58–73.

⁴² Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38, p. 16.

⁴³ In recent years, Fonagy's advocacy of manualized therapies has not diminished. However, he has allowed greater acknowledgement that, where manualization is not coupled with the potential to tailor treatment to the patient, this probably reduces efficacy because therapists prototype patients rather than recognize and respond to their specificity (see Chapter 7).

⁴⁴ Elsewhere Fonagy expressed worry that, if those sympathetic to psychoanalysis did not become engaged in trials, the modality would fail to have supporting evidence and lose reputation and access to public funding. And what trials were conducted would not be careful to pursue the complex work of identifying authentic indicators of success relevant to the modality. Fonagy, P. (2003). 'Psychoanalysis Today'. *World Psychiatry*, 2: 73–80.

Fourth, Fonagy argued that developments in theory needed to be tied more closely to developments in clinical technique, which could be evaluated to demonstrate or disconfirm the practical value of the theory. If psychoanalysts approach anxiety disorders, drug addiction, and postnatal depression differently on the basis of ideas about their respective psychological mechanisms, then this should be specified, and the distinct components of the interventions articulated to facilitate their empirical evaluation and the optimization of treatment practice. Fifth, psychoanalytic treatments should be revised to be shorter, with better specified goals to facilitate transparency and evaluation. And sixth, clinical progress should receive ongoing evaluation using reliable procedures, which also feed in to clinical supervision.⁴⁵ Of the six recommended changes to psychoanalytic practice, only this last would spare the Anna Freudian tradition, which had generally had a better history of evaluating clinical progress than other schools of psychoanalysis. On the five other grounds, Fonagy was directly calling for major alterations to the approach to psychoanalysis in which he had recently qualified.

The Anna Freud Centre retrospective study

In becoming research director for the Anna Freud Centre, Fonagy gained access to the Centre's case files for research purposes. It was expected that therapists associated with the Centre would write detailed weekly reports on each patient, and further reports every two months. These would belong to the Centre, for the purposes of research. Many of the case files had been indexed as part of a large project in the 1960s. The indexing work raised valuable questions about what should be the appropriate 'unit of psycho-analytic observation'. These questions led to penetrating studies of basic psychoanalytic concepts by Sandler and colleagues (see the Introduction), studies that were a formative influence for Fonagy.⁴⁶

By the time Fonagy took the role of research director, there were 763 cases available for research and, unusually, well characterized. Anna Freud and Dorothy Burlingham's sense of scientific values had led them to demand an unusual level of documentation. This included identification of symptoms such as conduct problems, anxiety, and depression. Cases at the Centre were also reporting against a quasi-interval system for coding mental ill health (see Chapter 7): category 1 (essentially normal), 2 (transient symptomatology or developmental strain), 3 (neurotic, with 'permanent regressions, fixations, and symptom formation'), 4 (atypical, distorted personality development), and 5 (destructive processes disrupting mental growth).⁴⁷ In the early 1990s, Mary Target began work on a PhD with Fonagy pursuing a retrospective study of patient records at the Anna Freud Centre to evaluate the effectiveness of the service. Target's characterization of Fonagy as a doctoral supervisor was as 'demanding', 'inspiring', and 'fun', though 'often infuriating' because 'much of the time I was left to get on with the work as best I could'.⁴⁸

⁴⁵ Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38, p. 16.

⁴⁶ Sandler, J. (1962). 'Research in Psycho-Analysis—The Hampstead Index as an Instrument of Psycho-Analytic Research'. *The International Journal of Psychoanalysis*, 43: 287–291; Fonagy, P. (2005). 'An Overview of Joseph Sandler's Key Contributions to Theoretical and Clinical Psychoanalysis'. *Psychoanalytic Inquiry*, 25(2): 120–147.

⁴⁷ Freud, A. (1962). 'Assessment of Childhood Disturbances'. *The Psychoanalytic Study of the Child*, 17: 149–158; Fonagy, P. and Target, M. (1994). 'Who is Helped by Child Psychoanalysis? A Sample Study of Disruptive Children, from the Anna Freud Centre Retrospective Investigation'. *Bulletin of the Anna Freud Centre*, 17: 291–315, p. 294.

⁴⁸ Target, M. (2003). 'The Anna Freud Centre: About the Directors'. Originally at <http://www.annafreudcentre.org/target.htm>. Accessed at: <http://web.archive.org/web/20060925220711/http://www.annafreudcentre.org/>

Fonagy and Target published their first reports from the study in 1994. At first sight, it appeared that psychoanalysis had very different outcomes for groups of child patients. Forty-six per cent of child patients, who initially presented with aggressive and disruptive behavioural problems, showed improvement in these symptoms by the end of analysis, compared with 73% of child patients presenting with symptoms of anxiety or depression. However, Fonagy and Target found that, if only analyses lasting at least three years were considered, the magnitude of change was identical. They interpreted these results as suggesting that 'it appears that psychoanalysis can bring about substantial improvements in children with disruptive disorders, but the challenge is to keep these children in analysis long enough for them to benefit'.⁴⁹ The researchers assessed factors such as whether there was a change in the therapist and whether there were regular meetings with the parents. These factors appeared to have no effect on the effectiveness of treatment of children showing aggressive and disruptive behaviour. Yet they predicted 20% of therapeutic outcome for children presenting with anxiety or depression.⁵⁰ Examining patient-level factors that predicted improvement, Fonagy and Target found that the age of the child did not affect the likelihood of improvement in their symptoms, but that older children were less likely to lose their diagnosis.⁵¹ Later meta-analytic findings from trials data by Pilling, Fonagy and colleagues would report that, in general, younger children obtain greater benefit from psychotherapy than older children by follow-up.⁵² However, a meta-analysis focused specifically on mentalization-based parenting interventions found the opposite, that interventions targeting parents of 6–12 year olds had greater effectiveness than those targeting parents of younger children.⁵³

Fonagy and Target acknowledged that there was considerable diversity of symptom profiles. So, for instance, within the 'aggressive and disruptive' group, there was actually much more improvement among children with oppositional defiant disorder (56%) than among those with a diagnosis of attention deficit hyperactivity disorder (36%) or conduct disorder (23%).⁵⁴ There was also heterogeneity in the anxiety and depression group, with poorer outcomes for depression.⁵⁵ Additionally, the distinction between 'aggressive and disruptive' behaviours and 'anxiety and depression' could be somewhat arbitrary. Many children displayed

target.htm; Target, M. (2018). '20/20 Hindsight: A 25-Year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy'. *Psychotherapy Research*, 28(1): 30–46.

⁴⁹ Fonagy, P. and Target, M. (1994). 'Who is Helped by Child Psychoanalysis? A Sample Study of Disruptive Children, from the Anna Freud Centre Retrospective Investigation'. *Bulletin of the Anna Freud Centre*, 17: 291–315, p. 302.

⁵⁰ Ibid. Good psychological functioning of the mother, and poor psychosocial adaptation of the child patient, at commencement of therapy were also positive predictors of improvements for children with anxiety and depression, but less so for children with aggressive or disruptive behavioural problems. Fonagy, P. and Target, M. (1996). 'Predictors of Outcome in Child Psychoanalysis: A Retrospective Study of 763 Cases at the Anna Freud Centre'. *Journal of the American Psychoanalytic Association*, 44: 27–77.

⁵¹ Target, M. and Fonagy, P. (1994). 'Efficacy of Psychoanalysis for Children with Emotional Disorders', *Journal of the American Academy of Child and Adolescent Psychiatry*, 33: 361–371, p. 367.

⁵² Pilling, S., Fonagy, P., Allison, E., Barnett, P., Campbell, C., Constantinou, M., ... and Kendall, T. (2020). 'Long-term outcomes of psychological interventions on children and young people's mental health: A systematic review and meta-analysis'. *PLoS One*, 15(11): e0236525.

⁵³ Lo, C. K. and Wong, S. Y. (2020). 'The effectiveness of parenting programs in regard to improving parental reflective functioning: a meta-analysis'. *Attachment & Human Development*, Early View.

⁵⁴ Fonagy, P. and Target, M. (1994). 'The Efficacy of Psychoanalysis for Children with Disruptive Disorders'. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33: 45–55, p. 45.

⁵⁵ Fonagy, P. and Target, M. (1996). 'Predictors of Outcome in Child Psychoanalysis: A Retrospective Study of 763 Cases at the Anna Freud Centre'. *Journal of the American Psychoanalytic Association*, 44: 27–77: 'We would argue that childhood depression may in some instances reflect a dysfunction of mental processes associated with the creation of self representation, particularly those associated with self-monitoring and self-evaluation' (pp. 54–55).

aspects of both profiles. An intriguing finding was that, among children presenting with aggressive and disruptive behavioural problems who remained in analysis for at least a year, 65% improved if there was an additional diagnosis of anxiety, whereas only 50% improved if there was no diagnosis of anxiety.⁵⁶ This suggested that anxiety was generally a *positive* prognostic indicator, a finding that Fonagy and Target attributed to the role of anxiety as a signal of psychological conflict, at least in older children, which could then be resolved in the analysis through support in reconciling different perspectives on the problem: ‘the absence of anxiety may then be an indication of a pervasive distortion of representations, or more likely a substantial inhibition of mental processes, to a point where incompatibility is no longer experienced.’⁵⁷ The finding that anxiety was a positive prognostic indicator has subsequently been supported by findings of a negative association between internalizing symptoms and impairment once general mental health (the p-factor, see Chapter 7) is taken into account.⁵⁸ Another related finding reported by Fonagy and Target was that children with mixed emotional and conduct disorder seemed particularly likely to benefit from four- or five-times weekly treatment, compared with once-a-week treatment. This was also the case for children with depression and those without a clear diagnostic profile. By contrast, children with milder forms of anxiety or with separation anxiety problems were just as likely to show improvements from once-a-week treatment.⁵⁹ A later meta-analytic study of child psychotherapy would indicate that moderate treatment intensity was more effective than high intensity across all disorders.⁶⁰

In a number of cases, the records showed that children had reported to their analyst abuse or neglect by their parents. Most clinicians took this seriously and approached the parents or worked with other professionals to try to resolve the problems. Nonetheless, a sad finding from the retrospective study was that some clinicians interpreted the children’s reports as fantasy. In a later follow-up of a proportion of the child patients, Target and Fonagy found that ‘at the time of follow-up these individuals could be seen to be reliving and being

⁵⁶ This proposal had already been suggested earlier by Conte, H. R., Plutchik, R., Picard, S. Karasu, T. B. and Vaccaro, E. (1988). ‘Self-Report Measures as Predictors of Psychotherapy Outcome’. *Comprehensive Psychiatry*, 29: 355–360.

⁵⁷ This was a complicated picture, however, and moderated by age. Fonagy and Target, ‘Predictors of Outcome’: ‘Variables specific to each age group emerged; for example, for children under 6 a history of mental illness in the mother predicted worse outcome, while sleep disorder, phobias, or significant medical history were associated with good outcome. Among latency children, in contrast, a history of maternal mental disorder was associated with better outcome. For adolescents, several variables related to parental mental health were significant predictors; for instance, an antisocial father or mother who had attempted suicide predicted worse outcome, but anxiety in the father was associated with improvement. Difficulties in peer relationships or disruptive behavior at school predicted poor outcome.’ (pp. 40–41). For later reflections on the potentially adaptive contribution of anxiety, see St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). ‘Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults’. *PLoS One*, 12(4): e0175381, pp. 19–20.

⁵⁸ Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... Moffitt, T. E. (2014). ‘The P Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders?’ *Clinical Psychological Science*, 2(2): 119–137; Lahey, B. B., Rathouz, P. J., Keenan, K., Stepp, S. D., Loeber, R., and Hipwell, A. E. (2015). ‘Criterion Validity of the General Factor of Psychopathology in a Prospective Study of Girls’. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 56(4): 415–422.

⁵⁹ Target, M. and Fonagy, P. (1994). ‘Efficacy of Psychoanalysis for Children with Emotional Disorders’. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33: 361–371, p. 368; Schachter, A. (2004). *The Adult Outcome of Child Psychoanalysis: A Long-Term Follow-Up Study*. Unpublished doctoral thesis, London: University College London.

⁶⁰ Pilling, S., Fonagy, P., Allison, E., Barnett, P., Campbell, C., Constantinou, M., ... and Kendall, T. (2020). ‘Long-term outcomes of psychological interventions on children and young people’s mental health: A systematic review and meta-analysis’. *PLoS One*, 15(11): e0236525.

preoccupied by the traumata, for example of intrusive memories or flashbacks that corresponded to actual experiences reported in the childhood file.⁶¹

The Anna Freud Centre retrospective study would be influential for the later thinking of Fonagy and colleagues in four regards in particular. First, Target and Fonagy identified pretend mode and psychic equivalence as forms of non-mentalizing through their work on the case files. This would form one of the most important bases for the development of their theory (see Chapter 5) and of Mentalization-Based Therapy by Fonagy and Bateman. Second, Anna Freud's quasi-interval scale for assessing mental ill health as a latent variable beneath symptom clusters seems to have been relevant background to work by Fonagy and colleagues on the p-factor in the 2010s (see Chapter 7). Third, and most directly, the study showed that psychoanalytic therapy could contribute to robust improvement in patients' symptoms over time, contrary to the claims of critics that suggested that psychoanalysis would either show no benefit or only had benefits that could not be documented. Fourth, children with anxiety problems improved just as well with once-a-week therapy as with four- or five-times-a-week therapy. This contributed to the interest of Fonagy and colleagues in lower-intensity therapies, and the delivery of mental health interventions for children and adolescents in primary care settings (see Chapter 8).

Transformations

The Anna Freud Centre retrospective study exemplified the willingness of Fonagy and colleagues to objectify and appraise the clinical efficacy of the Anna Freudian tradition of psychoanalytic intervention. Though Fonagy held the role of research director for the Centre, his stance was met by concern and scepticism by many of his fellow clinicians, who felt that his stance demonstrated insufficient appreciation for the richness of the Anna Freudian tradition in attempting to reduce its value to simplified outcome measures. Fonagy and his colleagues seemed to many at the Centre to be setting themselves up as independent of the community of Anna Freudians. He was also a controversial figure in the international psychoanalytic community, especially for his criticism of the concept of repression.⁶²

The issue of Fonagy's fidelity or infidelity to the Anna Freudian tradition would come to a head in 2002 with the retirement of Julia Fabricius as Director of the Centre. By this time, the Centre had a fourfold vision: conducting clinical training, analytic treatment, empirical research, and work on prevention. The Centre was comparatively rich, thanks to its estates in the heart of Hampstead. However, the institution was also financially overstretched in seeking to fulfil its different commitments.

Fonagy applied for the post of Director. In response, a petition was circulated to the board arguing against his appointment on the grounds that 'Peter Fonagy clearly stated at the AGM

⁶¹ Target, M. and Fonagy, P. (2003). 'Attachment Theory and Long-Term Psychoanalytic Outcome: Are Insecure Attachment Narratives Less Accurate?' in M. Leuzinger-Bohleber, A. U. Dreher, and J. Canestri (eds), *Pluralism and Unity? Methods of Research in Psychoanalysis*, London: International Psychoanalytical Association, pp. 149–167, p. 163. See also Schachter, A. and Target, M. (2009). 'The Adult Outcome of Child Psychoanalysis: The Anna Freud Centre Long Term Follow-Up Study', in Midgley N, Anderson J, Grainger E, Nescic-Vuckovic T (eds), *Child Psychotherapy and Research: New Approaches, Emerging Findings*, London: Routledge, pp. 144–156.

⁶² See Fonagy, P. (2000). 'Response'. *The International Journal of Psychoanalysis*, 81(2): 354–356. Leo Rangell, twice president of the International Psychoanalytical Association and the American Psychoanalytic Association, would describe Fonagy as having 'helped bring about' the 'current fragmentation' of psychoanalysis. Rangell, L. (2008). 'Reconciliation: The Continuing Role of Theory'. *Journal of the American Academy of Psychoanalysis and Dynamic Psychiatry*, 36(2): 217–233, p. 219. See also, Rangell, L. (2004). *My Life in Theory*, New York: Other Press.

that he does not think of himself as an Anna Freudian, and that, while prioritizing research, he would pursue cost-cutting measures that would discontinue clinical training and scale back the provision of treatment by the Centre. As a result, it was argued that 'we would go so far as to say that in his hands the identity and future of the Anna Freud Centre as an internationally renowned establishment dedicated to the Anna Freudian developmental perspective on child analysis would be destroyed.'⁶³ Thirty therapists and former therapists associated with the Centre signed the petition.⁶⁴ Fonagy withdrew his application. However, no other appropriate applicant could be found and Fonagy became Chief Executive the next year, with Mary Target and Linda Mayes as other members of the directorial team. Looking back on this moment of psychoanalytic history, Reeves has reflected that the petition:

now reads as a prescient though nostalgic letter, as the concerns raised in no way anticipated the sea-change the Centre would experience under Fonagy. It must also be said that no Anna Freudian child analyst or group of analysts came up with a clear vision to lead, protect, and fund the fourfold psychoanalytic clinic.⁶⁵

Nearly 20 years later, the Anna Freud Centre certainly looks a different institution from the one inherited by Fonagy. It should be immediately acknowledged that it is difficult to discern how much of this change was the result of the decisions and priorities of Fonagy, how much the decisions of others, and how much circumstance. Nonetheless, the net effect has been that the Centre's therapeutic, institutional, reputational, and financial resources have been almost wholly repurposed. While there are continuities in the strengths of all these elements, and some carry-overs such as parent-toddler groups,⁶⁶ the mission and scale of the Anna Freud Centre have shifted.⁶⁷ The Centre no longer offers long-term psychoanalytic therapy to children and adolescents. However, the Centre's Early Years Parenting Unit offers 18 months of two-days-a-week mentalization-based treatment to parents with a personality disorder with a child under five referred by child protection services.⁶⁸ Other clinical services are also

⁶³ Nagera, H. et al. (2002). 'Open Letter to the Chair of the Trustees.' The Anna Freud Centre. Accessed at: <http://www.thecjc.org/pdf/annafreud.PDF>

⁶⁴ White, K. and Schwartz, J. (2007). 'Attachment Here and Now: An Interview with Peter Fonagy'. *Attachment*, 1(1): 57–61.

⁶⁵ Reeves, J. K. (2017). 'About Losing and Being Lost without Anna Freud's 'Revolutionary Overhaul'. *Journal of the American Psychoanalytic Association*, 65(6): 1077–1101, p. 1079. The changes to the Anna Freud Centre under Fonagy's leadership can be regarded as foreshadowed by an earlier episode. In 1990, Fonagy joined the Executive Council of the International Psychoanalytic Association, serving as treasurer until 1995. Reports from this time reveal Fonagy's careful cost-effectiveness analysis of the Association's activities, and repeated confrontation of unpleasant truths about the viability of expenditure. This generated some controversy at times. However, by the end of Fonagy's tenure as treasurer, the Association was no longer on precarious footing, and in fact had started to build up reserves. Fonagy, P. (1990). 'Treasurer's column'. *International Psychoanalytical Association Newsletter*. Accessed at: <http://www.ipa.world/ipa/Images/PDFDocuments/IPA-News-Magazine/ipa-newsmag-1990-v22-i1.pdf>: 'We bear our own costs in mind... We aim to provide substance for the apocryphal remark "there is nothing quite as cost-effective as a good cost-effectiveness analysis."' (p. 9). Andrade de Azevedo, A. M. (1995). 'Secretary's Column'. *International Psychoanalytical Association Newsletter*. Accessed at: <http://www.ipa.world/ipa/Images/PDFDocuments/IPA-News-Magazine/ipa-newsmag-1995-v4-i2.pdf> 'It was considered remarkable that there had been hardly any over spending and indeed very little increase in costs for four years. Prof. Fonagy's report was received with applause' (p. 6).

⁶⁶ Zaphiriou Woods, M. and Pretorius, I. M. (2016). 'Observing, Playing and Supporting Development: Anna Freud's Toddler Groups Past and Present'. *Journal of Child Psychotherapy*, 42(2): 135–151.

⁶⁷ The most integrated and up-to-date account of the current work of the Anna Freud Centre is Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

⁶⁸ McLean, D. and Daum, M. (2017). 'The Use of Observation in Developing Parenting Capacity', in Clare Parkinson, Lucille Allain, and Helen Hingley-Jones (eds), *Observation in Health and Social Care: Applications for Learning, Research and Practice with Children and Adults*, London: Jessica Kingsley Publishers, pp. 158–176.

provided. For instance, the Parent–Infant Project offers a range of psychotherapeutic interventions focused on the parent–child relationship. The service is free for parents of infants living near the Centre who are struggling in the caregiving role. The Parent–Infant Project draws on elements of traditional parent–child psychotherapy, but also of mentalization-based therapy.⁶⁹ It delivers group-based outreach in primary care baby clinics, in collaboration with health visitors.⁷⁰ The Centre also offers eight-week, manualized groups, both for parents and for foster-carers, focused on supporting caregiver mentalization.⁷¹ Many of the clinical services delivered by the Centre have been adapted to virtual delivery during the COVID-19 pandemic, entailing both use of new technologies and adaptations of clinical technique.⁷²

Additionally, the Centre delivers innovative non-clinical services anticipated to have significant therapeutic benefit, and again these have been adapted for online delivery in the context of the pandemic.⁷³ For instance, the ‘Contact and Residence Disputes’ service takes a mentalization-based approach to helping families locked in chronic legal disputes about custody and contact for their child or children.⁷⁴ Such services signal the potential opening for a mentalization-based framework for professional practice across diverse contexts in which social understanding has difficulty forming or has broken down.

The training of child analysts has long gone. A clinical training in child psychotherapy remains, but the theoretical orientation of the course does not reflect an Anna Freudian perspective.⁷⁵ Besides this, psychoanalysis is in the background rather than the foreground at the Centre, though there remain courses such as the MSc Psychoanalytic Developmental Psychology. The Hampstead estates have been sold as part of generating the funds for a £41 million new building near King’s Cross Station in London, supported in part through private philanthropy and the patronage of the British royal family. In 2019, the building won the Grand Prix Design Award and was described by the judges as ‘a homely, light-filled environment that supports children dealing with mental health problems and who are

Accessed at: <https://www.annafreud.org/what-we-do/our-help-for-children-and-families/our-clinical-work-and-projects/our-work-with-under-fives/early-years-parenting-unit-eypu/>.

⁶⁹ There are several differences in the Anna Freud Centre model of parent–infant psychotherapy from the approaches adopted elsewhere. See Baradon, T., Biseo, M., Broughton, C., James, J., and Joyce, A. (2016). *The Practice of Psychoanalytic Parent–Infant Psychotherapy: Claiming the Baby* (2nd edn), London: Routledge.

⁷⁰ James, J. and Rosan, C. (2019). ‘Remodelling Baby Clinics: Opportunities to Support Parent–Baby Relationships’. *Journal of Health Visiting*, 7(8): 400–404.

⁷¹ Redfern, S., Wood, S., Lassri, D., Cirasola, A., West, G., Austerberry, C., ... Midgley, N. (2018). ‘The Reflective Fostering Programme: Background and Development of a New Approach’. *Adoption & Fostering*, 42(3): 234–248. See also Dueger, S. K. (2015). *An Attachment- and Mentalization-Focused Group: Experiences During One’s First Pregnancy*. Unpublished doctoral thesis, Chicago: Chicago University.

⁷² Ventura Wurman, T., Lee, T., Bateman, A., Fonagy, P., and Nolte, T. (2020). ‘Clinical management of common presentations of patients diagnosed with BPD during the COVID-19 pandemic: the contribution of the MBT framework’. *Counselling Psychology Quarterly*, Early View.

⁷³ Accessed at: <https://www.annafreud.org/mental-health-professionals/our-help-for-children-and-families/our-clinical-work-and-projects/parenting-and-multi-family-groups/family-ties-online-therapy-for-parents-in-conflict/>

⁷⁴ Accessed at: <https://www.annafreud.org/what-we-do/our-help-for-children-and-families/our-clinical-work-and-projects/our-work-with-families-involved-in-legal-proceedings/contact-and-residence-disputes/>. There is, as yet, no research evidence evaluating a mentalization-based approach compared with ‘mediation as usual’.

⁷⁵ Accessed at: <https://www.ucl.ac.uk/pals/study/pals-phd-and-doctorate-programmes/doctorate-psychotherapy-child-and-adolescent-psychoanalytic>. ‘The theoretical orientation of the training represents the thinking of the Independent School within the British Psychoanalytic movement. A Jungian pathway has been developed for those in Jungian analysis.’

excluded from the school system.’⁷⁶ In terms of the Centre’s focus, prevention has gained far greater prominence than in earlier incarnations, through the production of resources for schools, children and families, and helping professionals.⁷⁷ For instance, the Anna Freud Centre’s Service for Schools delivers training in a multi-family group approach, in which six to eight families receive mentalization-based support by a teacher at the school together with a psychological professional.⁷⁸ In the context of the COVID-19 pandemic, the Anna Freud Centre have issued a suite of prevention-focused resources and guidance for schools, children and families, and helping professionals.⁷⁹ For instance, they have published guidance on ‘Helping babies and young children under 5 through the coronavirus crisis’, which focuses on strategies parents and early years workers can use to help young children understand the disruption to their lives, allow them to continue to feel well looked after, and to identify escalating problems with their mental health.⁸⁰

On average, around 8,300 mental health professionals are trained a year at the Anna Freud Centre in mentalization-based approaches.⁸¹ As of 2018, over 5,000 mental health practitioners are affiliated with the Centre’s Learning Network, and over 6,500 teachers with the Centre’s Schools in Mind Network.⁸² In 2020, the Centre launched a new online training portfolio, which will have the advantage of making their courses much more readily available to practitioners outside of London, including internationally.⁸³

The Anna Freud Centre is also a major player in influencing national mental health policy, with a close relationship with the Department for Education. In 2014, the Centre became host to a new specialist school for children aged between 9 and 14 who had been excluded from mainstream education.⁸⁴ The scope of research at the Centre has also grown beyond recognition compared with 2003. The Centre receives over £2 million each year in grants for pursuing empirical research. It directly employs 63 researchers (out of 250 direct staff) and hosts 330 postgraduate students.⁸⁵

The Centre has a network of Young Champions—young people who have experience of using mental health services and contribute to research, governance, training and outreach.⁸⁶ On the Anna Freud Centre website is ‘On My Mind’, a section co-produced with

⁷⁶ Accessed at: <https://www.annafreud.org/insights/news/2019/11/double-award-win-for-our-new-home-the-kantor-centre-of-excellence/>.

⁷⁷ E.g. the ‘You’re never too young to talk’ animation and teacher toolkit, aimed at years 5 and 6 pupils: <https://www.annafreud.org/media/7228/tmh-parent-leaflet-final-all-approved-laid-out-for-web.pdf/>.

⁷⁸ Asen, E., Dawson, N., and McHugh, B. (2003). *Multiple Family Therapy: The Marlborough Model and its Wider Application*. London: Karnac Books.

⁷⁹ Accessed at: <https://www.annafreud.org/coronavirus-support/>

⁸⁰ Anna Freud National Centre for Children and Families (2020). Helping babies and young children under 5 through the coronavirus crisis, https://www.annafreud.org/media/11732/eyim_covidsupport_v1d2.pdf

⁸¹ The Anna Freud Centre (2018). *The Big Move: Annual Report and Financial Statements for the Year Ended 31 August 2018*. Accessed at: <https://www.annafreud.org/media/9623/trustees-annual-report-ye310818.pdf>.

⁸² Ibid.

⁸³ Accessed at: <https://www.annafreud.org/insights/news/2020/07/anna-freud-centre-launches-new-online-training-portfolio/>

⁸⁴ Fonagy, P. (2016). ‘Why Do Families Matter?’, Huffington Post, 17 February. Accessed at: http://www.huffingtonpost.co.uk/peter-fonagy/childrens-mental-health-families_b_9227210.html. In 2017, the school was rated as ‘outstanding’ by inspectors in every category. Accessed at: <http://www.thefamilyschoollondon.org/uploads/2014/03/The-Family-School-Ofsted-2017.pdf>.

⁸⁵ The Anna Freud Centre (2018). *The Big Move: Annual Report and Financial Statements for the Year Ended 31 August 2018*. Accessed at: <https://www.annafreud.org/media/9623/trustees-annual-report-ye310818.pdf>. The Anna Freud Centre (2019). ‘HRH The Duchess of Cambridge Launches Centre of Excellence to Drive Change for Child Mental Health’. Accessed at: <https://www.annafreud.org/insights/news/2019/05/hrh-the-duchess-of-cambridge-launches-centre-of-excellence-to-drive-change-for-child-mental-health>.

⁸⁶ The Anna Freud Centre (2020). ‘Young Champions’. Accessed at: <https://www.annafreud.org/on-my-mind/get-involved/young-champions>.

Young Champions, with information, signposting, and resources for children and young people to make sense of mental health and mental health services.⁸⁷ For instance, the Youth Wellbeing Directory provides a list of free mental health support organizations for young people, searchable by postcode or town. The website also has guidance for young people who are in a role that entails providing support for a friend or family member experiencing mental health difficulties.⁸⁸ Among the 'On My Mind' resources are self-care tools for young people, including support for activities such as creative writing, dance, conversations with friends, self-talk, time away from technology, and distraction techniques for when difficulties cannot otherwise be resolved.⁸⁹ These have been popular and well-used resources. Furthermore, the Centre has reported a 567% increase in use of these resources between March and April 2020, as the first COVID-19 lockdown commenced.⁹⁰

Despite the multifaceted nature of the transformation of the Anna Freud Centre over time, a few societal changes may be identified as having played an especially significant role at a structural level. These changes were identified and responded to early by Fonagy and colleagues, shaping the direction of their research and institution-building. They were: 1) an epistemic shift: the rise of evidence-based medicine; 2) a sociological shift: the prioritization of individual self-management; 3) a policy shift: the deintensification of health and welfare interventions for individuals with mental health needs; and 4) an epidemiological shift: the increasing identification of depression, anxiety, and self-harm among young people.

A first shift was the growth of evidence-based health care within Britain and greater appeal to the authority of outcome data in health policy decision-making and administration.⁹¹ From the late 1980s, pressure had begun to grow within the National Health Service for mental health treatments to supply evidence of their efficacy. On the one hand, data from randomized trials was important ammunition to justify or compete for public funding. On the other hand, collection of routine data regarding outcomes was called for in order to demonstrate continued short- and long-term cost-effectiveness, as well as the improvement of services over time.⁹²

Fonagy and Higgitt identified this trend in 1989, well ahead of the curve of their peers. They pleaded with the psychoanalytic community to respond to these oncoming pressures by the generation of evaluation research, and the development of robust and authentic quantitative measures for use in the collection of routine data.⁹³ This was part of the context for the Anna Freud Centre retrospective study. The findings were timely for decisions in Britain and internationally regarding the continued funding of psychoanalytic treatments by the state and medical insurance companies.⁹⁴ A further important development in the status

⁸⁷ Accessed at: <https://www.annafreud.org/on-my-mind>.

⁸⁸ Accessed at: <https://www.annafreud.org/on-my-mind/helping-someone-else>.

⁸⁹ Accessed at: <https://www.annafreud.org/on-my-mind/self-care>.

⁹⁰ <https://www.annafreud.org/insights/news/2020/04/the-duke-and-duchess-of-cambridge-support-every-mind-matters-and-chair-mental-health-roundtable-call>.

⁹¹ McLaughlin, K., Osborne, S. P., and Ferlie, E. (eds), (2002). *New Public Management: Current Trends and Future Prospects*, London: Routledge; Smith, R. and Rennie, D. (2014). 'Evidence-Based Medicine—An Oral History'. *JAMA*, 311(4): 365–367.

⁹² The question of how to capture young people's own perception of outcomes has remained a particular concern of the Anna Freud Centre over the years. See e.g. Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., and Wolpert, M. (2014). 'Measuring Mental Health and Wellbeing Outcomes for Children and Adolescents to Inform Practice and Policy: A Review of Child Self-Report Measures'. *Child and Adolescent Psychiatry and Mental Health*, 8(1): 14.

⁹³ Fonagy, P. and Higgitt, A. (1989). 'Evaluating the Performance of Departments of Psychotherapy'. *Psychoanalytic Psychotherapy*, 4(2): 121–153.

⁹⁴ Fonagy, P. and Target, M. (1994). 'Who is Helped by Child Psychoanalysis? A Sample Study of Disruptive Children, from the Anna Freud Centre Retrospective Investigation'. *Bulletin of the Anna Freud Centre*, 17: 291–315.

of Fonagy and his group at the Anna Freud Centre as a credible and trusted source of scientific knowledge on mental health in general, not just on psychoanalysis, was the completion of major systematic reviews of the evidence base of mental health interventions.⁹⁵ The first edition of *What Works for Whom?* was published in 1996 and reviewed over 2,000 studies.⁹⁶ A companion volume, focused on interventions for children and adolescents, was published in 2002.⁹⁷ In the wake of these publications, Fonagy and colleagues were commissioned in the early 2000s with large blocks of funding from the National Institute for Health and Clinical Excellence to develop clinical guidelines and outcome measures.⁹⁸ Fonagy was also appointed in 2002 to the Department of Health's Expert Group on Outcomes Measurement for mental health.

In Fonagy's view, the demand for data-driven mental health treatments was quite valid, a perspective that has become increasingly mainstream in recent years among mental health practitioners.⁹⁹ He felt that there would be benefits to patients of the movement towards evidence-based approaches to mental health treatment.¹⁰⁰ Yet, additionally, Fonagy acknowledged that in some respects 'the movement appears to be driven largely by financial considerations', as well as 'the motivation to impose social and political controls on the professional practitioner'.¹⁰¹ Rather than straightforwardly resist these pressures, Fonagy felt that the psychoanalytic community should respond to them through demonstrations of efficacy, cost-effectiveness, and manualization. This would then form the basis for arguments with policy-makers in terms that they would recognize. At a wider cultural level, he felt that in British society there was a growing expectation of demonstrable, short-term benefit from all activities and that psychoanalysis would have to adapt to this new reality.

A second societal shift, on which the work of Fonagy and colleagues has capitalized, has been sociological and cultural shifts towards a priority on individual self-management as the solution to social problems. Individuals are helped by their families, schools, workplaces, and other institutions to regulate and optimize their thoughts and feelings towards the end of coping with structural problems and economic precarity. This societal shift has sometimes

⁹⁵ Target, M. (2018). '20/20 Hindsight: A 25-Year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy'. *Psychotherapy Research*, 28(1): 30–46: 'Probably the central plank of the bridge towards our getting support for a large-scale RCT [randomized control trial] including child psychotherapy was broader work on systematic reviewing of the evidence base of interventions for childhood mental health problems' (p. 40). Fonagy states that 'To be a trusted source of knowledge is our key aim' at the Anna Freud Centre, in Kirby, T. (2019). 'Peter Fonagy—Battling the Enemy of Loneliness'. *The Lancet Psychiatry*, 6(12): 987.

⁹⁶ Roth, A. and Fonagy, P. (1996). *What Works for Whom: A critical Review of Psychotherapy Research* (1st edn). New York: Guilford Press.

⁹⁷ Fonagy, P., Target, M., Cottrell, D., Phillips, J., and Kurtz, Z. (2002). *What Works for Whom? A Critical Review of Treatments for Children and Adolescents*, New York: Guilford Press.

⁹⁸ Pilling, S. and Fonagy, P. (2012). 'Developing Clinical Guidelines for Children and Adolescents: Experience from the National Institute for Health and Clinical Excellence', in P. Sturmey and M. Hersen (eds), *Handbook of Evidence-based Practice in Clinical Psychology* (Vol. 1. *Child and Adolescent Disorders*), pp. 73–102. New York: Wiley.

⁹⁹ Fonagy, P. (2015). 'Peter Fonagy and the Undermining of Old Ideas on Personality Disorder'. Accessed at <https://www.escap.eu/research/peter-fonagy-and-the-undermining-of-old-ideas-on-personality-disorder/>: 'Clinicians used to be very sceptical and hostile, saying: "Hold on there, you are trying to undermine our status quo." Now a new, popular movement presents itself: a wide range of practitioners—from psychoanalytic to behavioural, to family and systemic oriented professionals—embrace the evidence-based way of working.'

¹⁰⁰ Murphy, M. and Fonagy, P. (2013). 'Mental Health Problems in Children and Young People', in *Our Children Deserve Better, Prevention Pays: Annual Report of the Chief Medical Officer* (Chapter 10), London: Department of Health: 'There is evidence that EBP [evidence-based practice] is statistically superior to usual care. Experimental work also demonstrates that the major benefit from EBP to child mental health services is in value, conceived of as the ratio of the outcome that matters to patients to the cost of delivering that outcome. Using EBP has been shown to reduce costs by up to 35% and duration of treatment by up to 43%' (p. 6).

¹⁰¹ Fonagy, P. (1999). 'Process and Outcome in Mental Health Care Delivery: A Model Approach to Treatment Evaluation'. *Bulletin of the Menninger Clinic*, 63(3): 288–304.

been referred to by sociologists as ‘neoliberalism’, to characterize a culture focused on individual responsibility and self-regulation, stoked by feelings of anxiety and emptiness should this self-regulation fail.¹⁰² The individual is encouraged to relate to their own life in an entrepreneurial way. This includes attention to thoughts and feelings, indeed one’s whole ‘personality’, as needing development and optimization.¹⁰³ Desires and attempts to image change are directed away from collective or policy solutions, treated as beyond reach, and instead turned inwards.

Christopher Bollas, from whom Fonagy originally took the ‘reflective’ in ‘reflective function’ (see Chapter 3), has recently observed that the thought of Fonagy and colleagues—like some of his own work—has reflected and responded to the increasing emphasis on individual self-management of thoughts and feelings.¹⁰⁴ Extrapolating on Bollas, it may be observed that, in a complex consumption-focused society, knowledge of the motivations and intentions implicated in one’s own thoughts and feelings is important for distinguishing utility, sifting wants from needs, selecting among available goods, and judging what price is worth paying.¹⁰⁵ Without this knowledge, we risk credulity, overextension, and flat-footedness. Furthermore, in a society in which the service sector accounts for the large majority of jobs, individuals have particular need of knowledge of the motivations and intentions implicated in the thoughts and feelings of others if they are to succeed in achieving sales, satisfying customers, receiving good ratings, and taking on debt in a strategic way. More generally, neoliberalism makes an individual’s ability to represent thoughts and feelings especially salient and important. Not only does this ability facilitate the optimal use of thoughts and feelings within an individual’s projects. It also facilitates a reduction in ways that thoughts and feelings may hinder these projects, or contribute more generally to social problems. Furthermore, the individual is held accountable for their thoughts and feelings, which reflects and responds to the wider perspective under neoliberalism in which individual responsibility or ‘ownership’ of their mental states and actions is emphasized, and in which the wider social, economic, and political context of mental states and actions is downplayed.¹⁰⁶

¹⁰² Binkley, S. (2011). ‘Psychological Life as Enterprise: Social Practice and the Government of Neo-Liberal Interiority’. *History of the Human Sciences*, 24(3): 83–102; Adams, G., Estrada-Villalta, S., Sullivan, D., and Markus, H. R. (2019). ‘The Psychology of Neoliberalism and the Neoliberalism of Psychology’. *Journal of Social Issues*, 75(1): 189–216.

¹⁰³ Scharff, C. (2016). ‘The Psychic Life of Neoliberalism: Mapping the Contours of Entrepreneurial Subjectivity’. *Theory, Culture & Society*, 33(6): 107–122; Teo, T. (2018). ‘Homo Neoliberalus: From Personality to Forms of Subjectivity’. *Theory & Psychology*, 28(5): 581–599; Gill, R. and Orgad, S. (2018). ‘The amazing Bounce-Backable Woman: Resilience and the Psychological Turn in Neoliberalism’. *Sociological Research Online*, 23(2): 477–495.

¹⁰⁴ Bollas, C. (2015). ‘Psychoanalysis in the Age of Bewilderment: On the Return of the Oppressed’. *International Journal of Psychoanalysis*, 96(3): 535–551. Bollas’s breakthrough paper was centrally concerned with individual self-management, and one of the detailed clinical cases is specifically focused on self-management in the workplace environment. Bollas, C. (1982). ‘On the Relation to the Self as an Object’. *The International Journal of Psychoanalysis*, 63: 347–359. ‘We are constantly engaged in acts of self management, from our choice of vocation to our choice of clothing, from our perception and facilitation of our needs to our management of our own personal realities to partially gratify those needs, from our recognition of, and planning for, holidays to our differing abilities to cognize and confront economic and familial realities.’ (p. 349).

¹⁰⁵ Cf. Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civali, C., Sartori, G., and Rumiati, R. (2008). ‘Mentalizing in Economic Decision-Making’. *Behavioural Brain Research*, 190(2): 218–223; Weiland, S., Hewig, J., Hecht, H., Mussel, P., and Miltner, W. H. (2012). ‘Neural Correlates of Fair Behavior in Interpersonal Bargaining’. *Social Neuroscience*, 7(5): 537–551.

¹⁰⁶ Sociologists have termed this societal shift ‘governmentality’, because the individual is asked to manage (govern) their own thoughts and feelings (mentality) in responding to social problems, rather than having recourse to collective or welfare responses to these problems. Schechter has adapted the term sociological ‘governmentality’ to refer to ‘governmentalization’, in characterizing the cultural alignment between a therapeutic focus on individual thoughts and feelings, on the one hand, and a policy focus on individual responsibility for coping with social problems, on the other. Schechter, K. (2014). *Illusions of a Future: Psychoanalysis and the Biopolitics of Desire*, Durham, NC, Duke University Press.

Individual skill at understanding social understanding and misunderstanding becomes all the more pertinent when institutions that might otherwise facilitate social understanding have been hollowed out and replaced with infrastructures generating competition, distrust, and feelings of alienation. Such skill is also all the more relevant in a multicultural society, in which common culture cannot be presumed upon. Over the years, Fonagy and colleagues have become increasingly critical of individual-centric modes of explanation, and acknowledged that societal values may have influenced Fonagy's own past thinking in this direction (see Chapter 9). They have also become increasingly concerned with alienation and culture, and the sociological conditions that undermine social understanding and trust.

A third trend over the past 30 years that has shaped the Anna Freud Centre has been the direction of public resources away from individualized and intensive health and welfare services. This can be regarded as the effect of two interrelated shifts in British health and welfare policy:

- Increasing moral weight put on egalitarianism, scalability, and population health.¹⁰⁷
- Increasing focus on cost-effectiveness, and indeed cheapness even in the absence of cost-effectiveness, with public services constructed as a potentially unfair burden on the taxpayer.¹⁰⁸

Again, Fonagy and colleagues caught this trend early, arguing in the Emanuel Miller lecture in 1992 for the need to 'envisage a mental health service with the capacity to meet the demand for help (let alone the underlying need) at the same time as being equitable, accessible and acceptable to all those who need it'.¹⁰⁹ Long-term psychoanalytic treatment of a few individuals seemed 'dissonant', he believed, both with crushing policy pressures to cut costs and with the democratic value of equitable access to mental health support for the huge numbers that could benefit from it.¹¹⁰ Fonagy argued that psychoanalysts should combat economic inequalities in access to mental health treatments: to do otherwise is to be morally complicit with the unequal distribution of wealth.¹¹¹ Short-term psychoanalytically inspired

¹⁰⁷ Mold, A., Clark, P., Millward, G., and Payling, D. (2019). *Placing the Public in Public Health in Post-War Britain, 1948–2012*, London: Palgrave. In Fonagy's view, this was a shift across many countries, but was especially intense in Europe. Fonagy, P. (2000). 'On the Relationship of Experimental Psychology and Psychoanalysis: Commentary by Peter Fonagy (London)'. *Neuropsychanalysis*, 2(2): 222–232, p. 227.

¹⁰⁸ Rao, A. S., Lemma, A., Fonagy, P., Sosnowska, M., Constantinou, M. P., Fijak-Koch, M., and Gelberg, G. (2019). 'Development of Dynamic Interpersonal Therapy in Complex Care (DITCC): A Pilot Study'. *Psychoanalytic Psychotherapy*, 33(2): 77–98. 'The Tavistock Adult Depression Study (TADS) (Taylor et al., 2012) has demonstrated the effectiveness of a psychoanalytic approach and the importance of follow-up data to consider the sustainability of change. However, the demand and capacity imbalance are increasingly putting pressures on services to cut down the length of the individual treatment' (p. 78).

¹⁰⁹ Fonagy, P., Steele, M., Steele, H., Higgitt, A., and Target, M. (1994). 'The Emanuel Miller Memorial Lecture 1992: The Theory and Practice of Resilience'. *Journal of Child Psychology and Psychiatry*, 35(2): 231–257, p. 231.

¹¹⁰ Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38, p. 8; Fonagy, P. and Bateman, A. (2009). 'A Brief History of Mentalisation-based Treatment and its Roots in Psychoanalytic Theory and Practice', in Brownescombe Heller, M., and Pollet, S. (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176: 'Preserving the key subtleties and respect for complexity of mind that psychoanalysis offers in its unparalleled richness, while at the same time enabling more than the few to have the benefit of its long-term value. This is the goal which we set ourselves on this journey that began 20 years ago' (p. 174). See also Kazdin, A. E. (2019). 'Annual Research Review: Expanding Mental Health Services through Novel Models of Intervention Delivery'. *Journal of Child Psychology and Psychiatry*, 60(4): 455–472.

¹¹¹ Jurist, E. L. (2010). 'Elliot Jurist Interviews Peter Fonagy'. *Psychoanalytic Psychology*, 27(1): 2–7: 'I think the social inequalities in health should concern us majorly as psychoanalysts. It's an embarrassment, because most psychoanalysts are in independent practice, and they are advantaged by the unequal distribution of wealth ... We should be looking at how we can work toward a more equal distribution of wealth, and in some ways try to prevent the extant inequalities from affecting future generations' (p. 5).

therapies could also offer a 'lifeboat' for psychoanalytic clinicians in the context of the preference among policy-makers for scalability and cost-effective interventions.¹¹²

Despite reservations from within the psychoanalytic community about the efficacy of empirical evaluation, Fonagy, Target, and colleagues were committed to the question of whether the effectiveness of psychoanalytic approaches could be demonstrated empirically at scale. Through the 1990s and 2000s, they attempted to find funding for a randomized control trial of child psychoanalytic psychotherapy, comparing it with alternative treatments. However, the response from reviewers was that it was not worth funding the study because it could never be cost-effective.¹¹³ Ultimately, they were able to find funding for a trial of 28 weeks of psychoanalytic psychotherapy for adolescents with depression. The results were published in 2017. The researchers found that psychoanalytic psychotherapy was no more effective than two other, briefer, treatments—as measured by patient self-report and clinical interview after 36 weeks, 52 weeks, and 84 weeks.¹¹⁴ All three treatments resulted in an average of 50% reduction in depression symptoms. The cost of delivering a Cognitive Behavioural Therapy intervention was £904.57 per patient, compared with £1,396.72 for psychoanalytic psychotherapy. However, as we saw earlier in the section on the Anna Freud Centre retrospective study, psychoanalytic treatment is comparatively ineffective for depression in children. Psychoanalytic psychotherapy was therefore tested with a condition and client group where Fonagy and colleagues anticipated that it would be less effective than with other conditions and client groups—presumably because this was where funding for a trial happened to be available. There may also have been 'sleeping effects' of the psychoanalytic psychotherapy, or benefits not accessible to self-report, as argued by Target.¹¹⁵ Evidence from another trial by Fonagy and colleagues of long-term psychoanalytic psychotherapy (60 sessions), compared with treatment as usual for adults with treatment-resistant depression, did find large differences at follow-up at 42 months (30.0% versus 4.4% partial remission of symptoms), suggesting that psychodynamic therapies can show their effectiveness in the longer term, though whether the findings from this latter trial were because of the longer follow-up, the specification of treatment-resistant depression, or the adult rather than adolescent patient group is unclear.¹¹⁶ According to the trial protocol, an economic evaluation was conducted

¹¹² Target, M. (2018). '20/20 Hindsight: A 25-Year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy'. *Psychotherapy Research*, 28(1): 30–46, p. 39. For a contrary position, arguing that the lifeboat offered by mentalization is too compromised to be worth reaching for, see House, R. (2012). 'General Practice Counselling amidst the 'Audit Culture': History, Dynamics and Subversion of/in the Hypermodern National Health Service'. *Psychodynamic Practice*, 18(1): 51–70.

¹¹³ Target, M. (2018). '20/20 Hindsight: A 25-Year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy'. *Psychotherapy Research*, 28(1): 30–46.

¹¹⁴ Goodyer, I. M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., ... and Senior, R. (2017). 'Cognitive Behavioural Therapy and Short-Term Psychoanalytic Psychotherapy Versus a Brief Psychosocial Intervention in Adolescents with Unipolar Major Depressive Disorder (IMPACT): A Multicentre, Pragmatic, Observer-Blind, Randomised Controlled Superiority Trial'. *The Lancet Psychiatry*, 4(2): 109–119; Goodyer, I. M., Reynolds, S., Barrett, B., Byford, S., Dubicka, B., Hill, J., ... and Senior, R. (2017). 'Cognitive-Behavioural Therapy and Short-Term Psychoanalytic Psychotherapy versus Brief Psychosocial Intervention in Adolescents with Unipolar Major Depression (IMPACT): A Multicentre, Pragmatic, Observer-Blind, Randomised Controlled Trial'. *Health Technology Assessment*, 21(12): 1.

¹¹⁵ Target, M. (2018). '20/20 Hindsight: A 25-Year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy'. *Psychotherapy Research*, 28(1): 30–46.

¹¹⁶ Fonagy, P., Rost, F., Carlyle, J. A., McPherson, S., Thomas, R., Pasco Fearon, R. M., ... and Taylor, D. (2015). 'Pragmatic Randomized Controlled Trial of Long-Term Psychoanalytic Psychotherapy for Treatment-Resistant Depression: The Tavistock Adult Depression Study (TADS)'. *World Psychiatry*, 14(3): 312–321. Treatment-as-usual consisted of various interventions, as directed by the referring practitioner, and could include referral to other specialist provisions. On age as a moderator of the effectiveness of interventions for depression, see Cuijpers, P., Karyotaki, E., Eckshtain, D. et al. (2020). 'Psychotherapy for Depression across Different Age Groups: A Systematic Review and Meta-Analysis'. *JAMA Psychiatry*, 77(7): 694–702.

of the cost-effectiveness of long-term adult psychoanalytic psychotherapy compared with treatment as usual. However, findings from this evaluation were not published.

A fourth trend to which Fonagy and the Anna Freud Centre have been especially responsive has been the changing situation of young people. It is a complicated picture, because the past 30 years have brought many forms of freedom for young people, especially those from socially and economically less deprived families. However, rates of depression, anxiety, and self-harm among young people have, at a population level, generally risen over the period.¹¹⁷ On the one hand, as Fonagy has observed, this likely reflects that as a society 'we have become better at expressing our thoughts, feelings, beliefs, wishes and desires and have passed this on to our children.'¹¹⁸ On the other hand, the epidemiology of adolescent mental illness likely partly reflects sociological changes.¹¹⁹ One contributing factor may have been increases in educational pressures in the context of frequent, high-stakes testing. Young people may have also experienced increasing worry about the consequences of failure in a social and economic context with fewer safety nets and the circulation of unrealizable images of what constitutes success.¹²⁰ Furthermore, Fonagy has described the pervasiveness and image-focused qualities of new media as important aspects of contemporary life, contributing to their potential influence on individual self-representations (see Chapter 9).

Fonagy's impression has been of a rise in despair among young people in the context of such changes, combined with a reduction in the availability and coherence of relevant community-based and third-sector supportive services.¹²¹

We see a growing mental health crisis across all groups. Trends highlight an increase in mental illness among some groups of children and young people, particularly emotional problems such as anxiety and depression... There can be little doubt that children and young people are experiencing new and multiple pressures in a demanding and fast-moving digitally enabled world.¹²²

¹¹⁷ Twenge, J. M., Joiner, T. E., Rogers, M. L., and Martin, G. N. (2018). 'Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates among US Adolescents after 2010 and Links to Increased New Media Screen Time'. *Clinical Psychological Science*, 6(1): 3–17; Patalay, P. and Gage, S. H. (2019). 'Changes in Millennial Adolescent Mental Health and Health-Related Behaviours over 10 Years: A Population Cohort Comparison Study'. *International Journal of Epidemiology*, 48(5): 1650–1664.

¹¹⁸ Fonagy, P. (2019). 'Mental Health is a Care We must Share'. *Guardian*, 13 October. Accessed at: <https://www.theguardian.com/society/2019/oct/13/isolation-not-social-media-cause-teenager-mental-ill-health>

¹¹⁹ Fonagy, P. (2016). 'We have Hard Choices to Make on Children's Mental Health'. *Huffington Post*, 10 October. Accessed at: http://www.huffingtonpost.co.uk/peter-fonagy/world-mental-health-day_b_12429138.html. See also, Fonagy, P. cited in Thomson, A. and Sylvester, R. (2020) 'Panic and Anxiety after Education is Plunged into Limbo'. *The Times*, 31 March: 'Between 20–27 per cent of female students now have common mental health disorders including depression, anxiety and phobias... There is greater pressure to achieve high grades and compete for jobs, from social media, financially and in what they think is expected of them as a woman, their goals and aspirations.'

¹²⁰ Cf. Blackman, S. and Rogers, R. (eds). (2017). *Youth Marginality in Britain: Contemporary Studies of Austerity*, Policy Press.

¹²¹ Fonagy, P. (2004). 'What Evidence for Evidence-based Prevention?' *Journal of Infant, Child, and Adolescent Psychotherapy*, 3(4): 419–443, p. 434; Fonagy, P. (2014). 'Evidence Submitted to the House of Commons Health Committee Children's and Adolescents' Mental Health and CAHMS'. Accessed at: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/health-committee/childrens-and-adolescent-mental-health-and-camhs/written/8009.html%23_edn1; Fonagy, P. and Allison, E. (2016). 'Commentary on Kernberg and Michels'. *Journal of the American Psychoanalytic Association*, 64(3): 495–500, p. 496.

¹²² Fonagy, P., Lenehan, C., and O'Sullivan, A. (2017). 'Foreword', in *Improving Mental Health Support for our Children and Young People*, London: Social Care Institute for Excellence, p. 2.

Payne and Fonagy have drawn attention to evidence that this mental health crisis is not evenly spread.¹²³ Low socio-economic status is associated with twice the rates of anxiety and depression as the rest of the British population. Gay, lesbian, bisexual, and queer young people are over two and a half times as likely to experience mental ill health. And girls show rising rates of anxiety, depression, and body dysmorphia as they move through adolescence, suggesting particular challenges for young women.¹²⁴

Fonagy as an individual, and the Anna Freud Centre as an institution, have been successful at presenting themselves as credible voices in the policy conversation about the future of children's mental health. On the one hand, these voices have been used to campaign for greater investment in the mental health needs of children and young people.¹²⁵ In evidence presented to the House of Commons Youth Select Committee in 2015, Fonagy strongly criticized a state of affairs in which children and young people received just over 6% of total national mental health expenditure. He regarded this as an injustice to the massive unmet need for mental health services for children and young people, and terrible foresight given the benefits of acting early in the life course to prevent or mitigate adverse developmental pathways.¹²⁶

Fonagy and the Anna Freud Centre have also increasingly had a role in the development of policy. Fonagy has chaired two national clinical guideline development groups, as well as several expert reference groups for the Department of Health. For example, the 2017 Green Paper on *Transforming Children and Young People's Mental Health Provision*, which put forward an increased role for schools in identifying and managing children's mental health problems, was based on a review co-authored by Fonagy. The work of Fonagy and colleagues in policy engagement in public mental health and preventative work will be discussed further in Chapter 9.¹²⁷

¹²³ Payne, C. and Fonagy, P. (2018). 'A Response to the Figures on the Mental Health of Children and Young People in England Released Today by NHS Digital'. Accessed at: <https://www.annafreud.org/insights/news/2018/11/a-response-to-the-figures-on-the-mental-health-of-children-and-young-people-in-england-released-today-by-nhs-digital>.

¹²⁴ Sadler K., Vizard T., Ford T., et al. (2018). *Mental Health of Children and Young People in England, 2017*. London: NHS Digital. Accessed at: <https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017>.

¹²⁵ Fonagy, P. and Pugh, K. (2017). 'Editorial: CAMHS Goes Mainstream' *Child and Adolescent Mental Health*, 22(1): 1–3.

¹²⁶ Youth Select Committee (2015). Oral evidence taken before the Youth Select Committee on 3 July. Accessed at: http://old.byc.org.uk/media/272544/youth_select_committee_03.07.15_pm.pdf

¹²⁷ Department of Health/Department of Education (2017). *Transforming Children and Young People's Mental Health Provision: A Green Paper*, London: HMSO.

3

Mentalizing in development

Introduction

Having situated Fonagy in biographical and institutional context in Chapters 1 and 2, we can now turn more directly to the development of the theory of mentalizing. This theory has undergone very substantial changes over time, and with these changes have come alterations in and additions to the meaning of the term ‘mentalization’. In their 2003 book, *Psychoanalytic Theories*, Fonagy and Target observed critically that it is quite characteristic of psychological theories to have a primary concept or two with a host of meanings. This concept then serves in part as a symbol of collective endeavour: ‘Many concepts referred to theoretically have multiple references, some pertaining to the developmental course, some to covert mental states and some to manifest presentation.’ Fonagy and Target highlighted that there are advantages to such a symbol, ‘particularly through enhancing a professional group identity by enabling members to believe that they share ideas, though ‘in the long term such fuzziness impedes progress, and scientific debate is degraded.’¹ The concept of ‘mentalization’ has also been subject to such a process, as Fonagy and colleagues have themselves increasingly acknowledged over the past decade.²

The term ‘mentalization’ is an umbrella term that has encompassed diverse meanings and investments. It offers the feeling of coherence that facilitates community building, while also serving as a switch and relay between research, health policy, and clinical practice, as well as between different clinical specialisms.³ At times, the term has been yoked to specific definitions or operationalization in psychological constructs, though these explicit definitions have seen very significant changes over time (see Chapter 4). At other times, ‘mentalization’ has functioned as a buzzword, used to mean little more than desirable psychological functioning,⁴ or a process that is in some way psychological.⁵ The concept of mentalization

¹ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 19.

² E.g. Fonagy, P. and Allison, E. (2012). ‘What is Mentalization? The Concept and its Foundations in Developmental Research’, in N. Midgley and I. Vrouva (eds), *Minding the Child: Mentalization-Based Interventions with Children, Young People and their Families*, pp. 11–34, Hove: Routledge, p. 32. This criticism of the concept was first made in a sustained way by Choi-Kain, L. W. and Gunderson, J. G. (2008). ‘Mentalization: Ontogeny, Assessment, and Application in the Treatment of Borderline Personality Disorder’. *American Journal of Psychiatry*, 165(9): 1127–1135: ‘Quickly adapted into psychiatric vernacular, the term mentalization has been used with variable meanings. The broad territory of the concept paradoxically contributes to its familiarity as well as to its ambiguity’ (pp. 1127–1128). More recently, see Tipple, R. (2017). ‘Thinking versus Mentalization’. *Art Therapy Online*, 8(2).

³ On the crafting of an idea and ideal of a knowable object of study as a guarantee to sustain coalitions of practice, identification, and research, see Wiegman, R. (2012). *Object Lessons*, Durham, NC: Duke University Press.

⁴ On buzzwords, see Bensaude Vincent, B. (2014). ‘The Politics of Buzzwords at the Interface of Technoscience, Market and Society: The Case of “Public Engagement in Science”’. *Public Understanding of Science*, 23(3): 238–253.

⁵ Slade, among others, has observed that the term ‘mentalization’ seems to operate with narrow and broader meanings: Slade, A. (2005). ‘Parental Reflective Functioning: An Introduction’. *Attachment & Human Development*, 7(3): 269–281. To take one example: Fonagy and Adshead suggest that ‘classical behaviour therapy involves mentalisation. When patients are taught new responses to fear-inducing stimuli, changes occur in the ventral prefrontal cortex and amygdala, which are involved in fear reactions.’ The capaciousness of the term ‘mentalization’ is visible here, entailing only that mental states are identified and implicated. This makes

appears as an inspiring ideal, and one with multiple possible uses. However, this picture has been achieved in part by punching holes in the concept's ability to hold determinate meaning and facilitate clear communication across domains. This predicament has made discrepant research findings difficult to interpret, because researchers may well be measuring different phenomena but nonetheless calling them 'mentalization'.⁶ A few of Fonagy's collaborators have even argued that the term 'mentalization' should not be used with patients, given its untamed panoply of semantic and theoretical investments.⁷ However, this is not the majority view.

To unpack the different meanings the concept has gained over time, it is necessary to take a developmental approach. The focus of the chapter will be work by Fonagy and colleagues addressed to the field of developmental psychology in the 1990s, though later theory and research will be mentioned when there have been continuities. We will begin by considering the emergence of the concept of mentalization, the problems it was introduced to address, the theoretical perspective it encapsulated, and the clinical implications that stemmed from this perspective. We will then examine the development of the reflective functioning scale resulting from Fonagy's work with Miriam and Howard Steele, a very important development for work on mentalizing, as well as for the broader fields of developmental and clinical psychology. The chapter will close with an analysis of some ambiguities and problems in the use of the concept of disorganized attachment by Fonagy and colleagues.

Origins of the mentalization concept

As discussed in Chapter 1, the idea of a 'borderline' profile of mental illness had emerged in the twentieth century and gained prominence from entering the DSM-III in 1980 as one of the personality disorders. In the 1980s and 1990s, the wider clinical community regarded borderline personality disorder (BPD) as untreatable, even as a growing proportion of patients in NHS psychotherapy services were identified as having this profile.⁸ However, clinicians at the Anna Freud Centre regarded these cases as requiring an adjustment of clinical

mentalization relevant to every psychological modality and specialism. Fonagy, P. and Adshead, G. (2012). 'How Mentalisation Changes the Mind'. *Advances in Psychiatric Treatment*, 18(5): 353–362, p. 35. But surely to claim that classical conditioning, effective with rats and artificial intelligences, implies mentalization is difficult to reconcile with Bateman and Fonagy's majestic claim that 'mentalising is a uniquely human capacity—it can be seen as what defines humanity': Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 4. This statement might be interpreted as reflecting Fonagy's earlier claim that 'No animal, not even the most intelligent of non-human primates, can discern the difference between the act of a conspecific due to serendipity and one rooted in intention, wish, belief or desire.' Fonagy, P. (2012). 'Mentalization and Attachment: The Implication for Community Based Therapies'. Paper presented at the Community of Communities 10th Annual Forum, March. London. Accessed at: <https://www.slideshare.net/raffaelebarone/mentalization-and-attachment-the-implication-for-community-based-therapies>.

⁶ E.g. Ballespi, S., Vives, J., Debbané, M., Sharp, C., and Barrantes-Vidal, N. (2018). Beyond Diagnosis: Mentalization and Mental Health from a Transdiagnostic Point of View in Adolescents from Non-Clinical Population. *Psychiatry Research*, 270: 755–763.

⁷ Reported in Keaveny, E., Midgley, N., Asen, E., Bevington, D., Fearon, P., Fonagy, P., ... Wood, S. (2012). 'Minding the Family Mind: The Development and Evaluation of Mentalization Based Treatment for Families at the Anna Freud Centre in London', in N. Midgley and I. Vrouva (eds), *Minding the Child: Mentalization-Based Interventions with Children, Young People and their Families*, London: Routledge, pp. 98–112, p. 102.

⁸ For later discussion of this attitude, see Chiesa, M., Fonagy, P., Bateman, A. W., and Mace, C. (2009). 'Psychiatric Morbidity and Treatment Pathway Outcomes of Patients Presenting to Specialist NHS Psychodynamic Psychotherapy Services: Results from a Multi-Centre Study'. *Psychology and Psychotherapy: Theory, Research and Practice*, 82(1): 83–98.

technique to emphasize understanding and support. Looking back, Bateman and Fonagy observed in 2003 that the diagnostic category of BPD:

has evoked intense theorizing among psychoanalysts, behaviourists, psychiatrists, and others, and, perhaps because of its clinical difficulty and variability, that it represents a battlefield on which many of the controversies and schisms are played out.⁹

In the DSM-III, 'borderline' was considered to be a personality disorder because it signalled widespread difficulties an individual faced in understanding and regulating their 'self' in relation to others. In the 1980s, it was assumed that pervasive psychosocial symptoms had to be located in the 'personality' of the patient.¹⁰ Running somewhat contrary to this focus on individual personality was growing concern in the period with the socially situated self in developmental psychology, with the solidification of research concerned with social cognition. This represented important acknowledgement of social and developmental contributions to the mind of the individual (albeit with only somewhat greater acknowledgement of the role of culture).¹¹ Within both British and American developmental psychology in the 1980s, this shift was fed by several currents.

One was attachment theory: in the third book of his *Attachment and Loss* trilogy, Bowlby had reinterpreted individual differences in attachment as reflecting variation in the filtering and processing of social information about the availability of attachment figures.¹² Research in American developmental science had also started to explore empirically how children develop an understanding of mental phenomena, conceptualized as the child's 'theory of mind'.¹³ These perspectives helped scaffold empirical research by scholars such as Bretherton and Dunn on communication within families, which explored how children and parents spoke about feeling states and intentions.¹⁴ A further current was the growing maturity of empirical research on autism. Researchers such as Baron-Cohen had drawn on the idea of 'theory of mind' to suggest that autism reflected differences and/or deficits in the capacity of an individual to understand the mental states of other people.¹⁵ Research on social cognition was diverse, but had in common that it questioned the assumption that human experience of self and others is prewired, universal, and equally accessible. Instead, it was proposed that such experience depends upon socio-emotional, communicative, and neurodevelopmental supports, which vary substantially between individuals.

⁹ Bateman, A. W., and Fonagy, P. (2003). 'The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder'. *Bulletin of the Menninger Clinic*, 67: 187–211, p. 188

¹⁰ Cf. Fonagy, P. (2007). 'Personality Disorder'. *Journal of Mental Health*, 16(1): 1–4: 'We suspect now that it is a considerable oversimplification to assume that a pre-existing characterological basis to psychopathology can provide a breeding ground for psychiatric disorder.' (1) Pickersgill, M. (2013). 'How Personality Became Treatable: The Mutual Constitution of Clinical Knowledge and Mental Health Law'. *Social Studies of Science*, 43(1): 30–53.

¹¹ Jovanovic, G. Alolio-Näcke, L., and Ratner, C. (eds) (2018). *The Challenges of Cultural Psychology: Historical Legacies and Future Responsibilities*, London: Routledge.

¹² Bowlby, J. (1980). *Loss*, New York: Basic Books; Bretherton, I. (1987). 'New Perspectives on Attachment Relations: Security, Communication and Internal Working Models', in J. Osofsky (ed.), *Handbook of Infant Development*, New York: John Wiley, pp. 1016–1100.

¹³ Wellman, H. M. and Estes, D. (1986). 'Early Understanding of Mental Entities: A Reexamination of Childhood Realism'. *Child Development*, 57: 910–923. See also Flavell, J. H. (2004). 'Theory-of-Mind Development: Retrospect and Prospect'. *Merrill-Palmer Quarterly*, 50(3): 274–290.

¹⁴ Bretherton, I., McNew, S., and Beeghley-Smith, M. (1981). 'Early Person Knowledge as Expressed in Gestural and Verbal Communication: When do Infants Acquire a "Theory Of Mind"?', in M. E. Lamb and L. R. Sherrod (eds), *Infant Social Cognition*, Hillsdale, NJ: Erlbaum; Dunn, J., Bretherton, I., and Munn, P. (1987). 'Conversations about Feeling States between Mothers and their Young Children'. *Developmental Psychology*, 23(1): 132–139.

¹⁵ Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). 'Does the Autistic Child have a "Theory of Mind"?'. *Cognition*, 21(1): 37–46.

Reflecting on clinical cases such as Mr S (Chapter 1), in the late 1980s and early 1990s, Fonagy found a 'point of contact' between the developmental theory of research in social cognition and the profile of symptoms characterized as BPD.¹⁶ His proposal was that the difficulties with identity, mood, empathy, and social relationships associated with BPD could be conceptualized as expressions of a single underlying problem: disruption in the capacity to process and reconsider information about mental states. In therapy, Fonagy's patients with BPD described childhoods characterized by a lack of parental empathy, consistency, and/or safety. Fonagy knew, of course, that the recollections of adults in therapy needed to be treated with caution. He also knew that children's experiences of their parents may be inflected by their hopes and worries, as well as by the adult's actual behaviour.¹⁷

Nonetheless, he extrapolated that experiences that would make it difficult for children to tolerate thinking about the mental states of their parents would hinder the development of social understanding more generally. The basic underpinning mechanism, Fonagy anticipated, was the classic psychoanalytic premise of the 'pleasure principle'. This principle supposes that children will avoid displeasure wherever possible. Fonagy hypothesized that children will turn away from thinking about mental states if these were sources of unpleasure or outright pain.¹⁸ Reflecting on the case of Mr S., Fonagy proposed that the most direct cause would be when parents were 'unloving and cruel', because this would make 'the contemplation of the contents of the mind of the object unbearable'.¹⁹ His clinical work and review of existing retrospective research on BPD led him to conclude that sexual and physical abuse would make an especially potent contribution.²⁰ However, he also suggested that lack of parental empathy, consistency, and the provision of safety might contribute through various means. For instance, he argued that a 'totally vacuous relationship' might be expected to have the same effect as an abusive or hostile one, in making mental states a source of unpleasure.²¹

¹⁶ Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 86. This point of contact was further facilitated by a symposium held to facilitate dialogue between developmental scientists and clinicians at the Anna Freud Centre. Moran, G. S. (1990). 'Report on the One-Day Symposium between European Developmental Scientists and Anna Freud Centre Staff'. *Bulletin of the Anna Freud Centre*, 13(3): 167–168.

¹⁷ Cf. Fonagy, P. (2011). 'Discussion of Juan Pablo Jimenez's Paper, "A Fundamental Dilemma of Psychoanalytic Technique. Reflections on the Analysis of a Perverse Paranoid Patient"', in J. P. Jimenez and R. Muguillansky (eds), *Clinical and Theoretical Aspects of Perversion*, London: Karnac Books, pp. 63–76: 'Of course, the malevolent thoughts of the adult may not be real, but may be anticipated by the child, giving rise to a phantasy that may be the product of neurotic compromise, as when the Oedipal child imagined parental retribution' (pp. 64–65).

¹⁸ Fonagy, P. (1991). 'Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient'. *The International Journal of Psychoanalysis*, 72: 639–656: 'In the case of borderline functioning, I believe the deficit to be self-imposed and partial. It is brought about by a defensive disavowal of the mental existence (in terms of psychic functioning) of the object. Such disavowal is undertaken in the face of an anticipation of unbearable psychic pain' (p. 651).

¹⁹ *Ibid.* 650.

²⁰ Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients'. *Bulletin of the Anna Freud Centre*, 12: 91–115, pp. 107–108. In fact, later research would show that physical abuse was less predictive of personality disorder than other forms of maltreatment. Cirasola, A., Hillman, S., Fonagy, P., and Chiesa, M. (2017). 'Mapping the Road from Childhood Adversity to Personality Disorder: The Role of Unresolved States of Mind'. *Personality and Mental Health*, 11(2): 77–90, Table 2.

²¹ Fonagy, P. (1996). 'The Significance of the Development of Metacognitive Control over Mental Representations in Parenting and Infant Development'. *Journal of Clinical Psychoanalysis*, 5(1): 67–86: 'In cases of an abusive, hostile, or simply totally vacuous relationship with the caregiver, the infant may deliberately turn away from the mentalizing object because the contemplation of the object's mind is overwhelming as it harbors frankly hostile intentions toward the infant's self' (p. 84). For criticism of the necessity of such speculative attributions by the infant to the caregiver, see Fuchs, T. and De Jaegher, H. (2009). 'Enactive Intersubjectivity: Participatory Sense-Making and Mutual Incorporation'. *Phenomenology and the Cognitive Sciences*, 8(4): 465–486.

During the 1990s, Fonagy was not sure whether trauma was the cause of the inhibition in social cognition, or a correlate; and whether the inhibition could occur in the absence of discrete trauma, as in the case of a 'vacuous' relationship. Poor specification of the concept of trauma also meant that he sometimes lost track of these questions altogether, especially when trauma was itself defined as the absence of mentalization.²² Nonetheless, in general, in this period he held that trauma in the early child-parent relationship would especially predispose a turning away from attention to mental states throughout development.²³ However, he accepted that other aversive experiences of caregiving could have the same effect:

Inhibition of mental processes normally arises in the context of intense trauma, but this does not necessarily involve parental disturbance. Such inhibition does, however, require an internal experience that is sufficiently intense to make a permanent distortion of a basic mental capacity. This was the situation in the case which follows. Joan had had a bone disease from the age of 5 weeks which persisted throughout the first two years of her life. She suffered from pain which spoiled all the normal, pleasurable, physical interactions with her mother; being cuddled, washed, dressed, and fed all hurt. In addition her mother had to help in painful and frightening medical procedures. This resulted in a deeply ingrained view of mother as attacker and of herself as hurt, damaged, and deprived. The mother in turn felt helpless and inadequate and severe distortions of object representations were inevitable. Because the early mother child interactions were painful rather than pleasurable, they gave rise to severe inhibition and stunting of mental processes.²⁴

Influenced by the tendency of psychoanalytic theories to speculatively assign the origin of particular mental health problems to a developmental stage of early life, Fonagy gave particular prominence to the second year of life. He argued that children of this age would have the cognitive maturity to be capable of inhibiting attention to mental states, while still lacking more mature and differentiated strategies for responding to abusive, hostile, or 'vacuous' other minds.²⁵

On the basis of this developmental perspective, Fonagy proposed that an inhibition of the capacity to represent mental states, perhaps initiated in toddlerhood, could account for the symptoms of BPD: disruptions in an individual's sense of their identity, their identification and regulation of feelings, their capacity to empathize with others, and generally their capacity to engage in satisfying social relationships. To resolve these problems, Fonagy proposed that the capacity to represent mental states needed to be reignited: 'Only the gradual

²² E.g. Fonagy, P. (2003). 'Clinical Implications of Attachment and Mentalization: Efforts to Preserve the Mind in Contemporary Treatment. Epilogue'. *Bulletin of the Menninger Clinic*, 67: 271–280: 'Trauma may be defined subjectively as an experience that is intolerable to entertain: unmentalizable' (p. 272).

²³ The status of trauma in Fonagy's early thinking as probably one sufficient but not necessary condition for problems with the capacity to conceptualize mental states is already visible in Fonagy, 'On Tolerating Mental States'. This position would later be qualified: Fonagy, P., Bateman, A. W., and Luyten, P. (2012). 'Introduction and Overview', in *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 3–42: 'General characteristics of family function rather than maltreatment per se generate vulnerability to loss of mentalisation under stress' (pp. 15–16); 'the mentalisation-based treatment model does not attribute a central role to trauma ... In our view the effect of trauma is most likely to be felt as part of a more general failure to consider the child's perspective' (p. 16).

²⁴ Fonagy, P., Edgumbe, R., Target, M., and Miller, J. (1999). *Contemporary Psychodynamic Child Therapy: Theory and Technique*, London: The Anna Freud Centre and University College London, pp. 46–47.

²⁵ Fonagy, P. (1991). 'Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient'. *The International Journal of Psychoanalysis*, 72: 639–656, p. 654.

elaboration of mental representations of thoughts and feelings will free them from the crushing grip that the concrete experience of reality imposes upon the human mind.²⁶

Where the mental states of self and other can be drawn upon as a source of clear, specific, relevant, and transferable information, this permits an individual to make gradual and constant adjustments to their expectations about social interactions. The blockage of this process was likewise regarded by Fonagy as the basis for the extreme stances of patients with BPD, as well as contributing to a sense of grievance with a world that will not align with their expectations.²⁷ Fonagy held that blocks on thinking about mental states could also account for other symptoms of BPD. For example, the feelings of disintegration reported by Fonagy's patients could be understood in this perspective as reflecting difficulties in integrating self-relevant personal and social information, which would usually stabilize a sense of self. The feelings of emptiness reported by the patients could be understood as reflecting blocks in what would otherwise be the vivid sensations from experiences of self and other minds, and the shallowness with which relationships are therefore experienced: 'The loss of the underlying experience of shared consciousness makes the whole world appear flat, meaningless and isolating.'²⁸ Without shared consciousness, there is a poverty of sources of value, given that values are inherently social.

Fonagy cited his debt to research on social cognition, and especially the work of Baron-Cohen and Dunn. These researchers had demonstrated that the capacity for social understanding was a developmental achievement. However, Fonagy felt that Baron-Cohen tended to depict the child as an independent cognitive agent, alone in attempting to know the world. This neglected the centrality of child-caregiver interaction and emotional relationship, as an important environment within which children learn what it means to recognize a feeling, thought, or intention within themselves or in others. These early relationships also shape children's feelings about such recognition, including whether it will likely offer benefits or whether punishments will ensue.²⁹ In contrast to Baron-Cohen, the work of Dunn was closer

²⁶ Fonagy, P. (1995). 'Playing with Reality: The Development of Psychic Reality and its Malfunction in Borderline Personalities'. *The International Journal of Psychoanalysis*, 76(1): 39–44, p. 43.

²⁷ Fonagy, P. (1993). 'Psychoanalytic and Empirical Approaches to Developmental Psychopathology: An Object-Relations Perspective'. *Journal of the American Psychoanalytic Association*, 41: 245–260: the capacity to think about mental states 'equips the individual with a sort of ballast, a self-righting capacity where working models may become the object of review and change. Such gradual and constant adjustments give rise to an internal world where the behavior of objects can be experienced as predictable and stable, where the need for splitting of incoherent mental representations of the other is minimized and new experiences can be readily integrated with past internal representations' (p. 256).

²⁸ Fonagy, P. and Target, M. (2007). 'Playing with Reality: IV. A Theory of External Reality Rooted in Intersubjectivity'. *The International Journal of Psychoanalysis*, 88: 917–937, p. 921. See also Fonagy, P. (2000). 'Attachment and Borderline Personality Disorder'. *Journal of the American Psychoanalytic Association*, 48(4): 1129–1146: 'The emptiness is a direct consequence of the absence of secondary representations of self states, certainly at the conscious level, and of the shallowness with which other people and relationships are experienced... Emptiness (and in extreme cases a sense of dissociation) is the best description such individuals can give of the absence of meaning that the failure of mentalization creates' (p. 1140). Fonagy and colleagues may now regard this emptiness as expressing aspects of the primary unconscious—see Chapter 6. This would seem to be the position of Grotstein, J. S. (1984). 'A Proposed Revision of the Psychoanalytic Concept of Primitive Mental States, Part II: The Borderline Syndrome-Section 2: The Phenomenology of the Borderline Syndrome'. *Contemporary Psychoanalysis*, 20(1): 77–119; Grotstein, J. S. (1990). 'Nothingness, Meaninglessness, Chaos, and the "Black Hole" I—The Importance of Nothingness, Meaninglessness, and Chaos in Psychoanalysis'. *Contemporary Psychoanalysis*, 26: 257–290. However, unlike Fonagy and colleagues, Grotstein differentiates between emptiness as unintegration—the uncommitted potential for meaning-making—and emptiness as disintegration. The distinction is between a container with nothing yet in it and a broken container in which meaning has drained away.

²⁹ Implicit in Fonagy's early remarks on Baron-Cohen, this was brought out explicitly a few years later in Target, M. and Fonagy, P. (1996). 'Playing with Reality: II. The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479: 'Our position is at odds with the majority of developmental psychologists (e.g. Leslie, 1987; Baron-Cohen, 1995). Current psychological theories stress mainly cognitive precursors of theory of mind. The child is seen as an isolated processor of information, engaged in the

to Fonagy, because Dunn had suggested the importance of family interaction for scaffolding children's understanding of the i) origin, ii) location, and iii) functioning of mental states. However, Dunn had also considered children's iv) understanding of social and cultural conventions under the rubric of social cognition, and this more sociological set of meanings was not of such interest to Fonagy, at least in the 1990s (see Chapter 9).³⁰ He was much more concerned with the individual's capacity to attend specifically to their own mind and the mind of concrete others.

The term 'mentalization' was introduced to describe social cognition regarding the origin, location, and functioning of mental states. A desire on the part of Fonagy to circumscribe the object of concern was understandable—especially in retrospect, given that the concept of 'mentalization' has ballooned dramatically. However, the decision to cut understanding of social and cultural conventions from mentalizing would be incredibly important in circumscribing the scope of causal explanations and theoretical concern, and for kicking into the long grass questions of context and culture that would haunt the paradigm over decades, as Fonagy and colleagues have recently acknowledged (see Chapter 9). Above all, the role of social and cultural conventions and processes in *structuring* the origin, location, and functioning of mental states—and who is attributed mental states at all—was ejected from view. In Lacanian terms, Fonagy's decision to exclude concern for understanding rules and conventions from mentalizing limited the theory to concern with the felt experience of self and others (what Lacan terms 'the imaginary'), neglecting the complex and conflicted cultural context that structures our felt experience of self and others ('the symbolic').³¹ There was no intrinsic reason why a concern with individual perceptions of minds could not have been integrated with attention to social and cultural conventions. However, at the time, it was not, with the result that culture had to be rediscovered in the 2010s (see Chapter 9).

construction of a theory of mind from first principles, accumulating representations of the world through observation. From a psychoanalytic viewpoint this is a barren picture that ignores the central role of the child's emotional relationship with the parents or other caregivers in fostering the capacity to understand interactions in terms of mental states . . . We may assume, as do most cognitivists working in this area, that the development of a theory of mind is canalized (Waddington, 1966) or prepared, but the canal is dug not by biology, but by infant-parent interaction' (p. 461).

³⁰ Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients'. *Bulletin of the Anna Freud Centre*, 12: 91–115: 'Dunn (1988) considers four aspects of this development: 1) understanding others' feelings, 2) understanding others' intentions, 3) understanding rules and conventions and 4) understanding that others have minds' (p. 91).

Components of social cognition for Dunn (1988)	Components of mentalizing for Fonagy (1989)
Understanding the origin of mental states	Understanding the origin of mental states
Understanding the location of mental states	Understanding the location of mental states
Understanding the functioning of mental states	Understanding the functioning of mental states
Understanding of social and cultural rules and conventions	

³¹ Lacan, J. ([1953–1954] 1988). *Freud's Papers on Technique*, trans. John Forrester, NY: Norton. Among other consequences, Fonagy's evacuation of the symbolic rather depoliticized the concept of mentalizing. So, for instance, when Fonagy conceptualized conduct problems as caused by individual non-mentalizing (see Chapter 6), he did so without adequate attention to the role of inequalities in shaping conventions for the attribution of mental states and the role of institutionalized non-mentalizing. There was no intrinsic reason why a concern with individual perceptions of minds could not have been integrated with attention to social and cultural conventions. However, at the time, it was not, with the result that this concern has had to be rediscovered in the 2010s (see Chapter 9).

On his first use of the word in 1989, Fonagy described it as a contraction of the phrase ‘capacity to conceive of mental states’.³² The term presented itself in part because of an existing usage in Paris.³³ Marty had been the president of the Parisian Psychoanalytic Society when Ivan Fónagy moved to the city and began training as a psychoanalyst. In 1972, Marty was co-founder of a clinic specializing in the treatment of ‘psychosomatic’ symptoms. The École Psychosomatique de Paris came to conceptualize these symptoms as expressing unprocessed mental states, which had been defensively blocked from achieving adequate and flexible representation.³⁴ The capacity to form such representations was termed by Marty and colleagues as ‘mentalization’. Presumably conversations between Peter and his father contributed to an awareness of these developments even before they reached their full and widely known statement in Marty’s 1990 book, *Mentalisation et Psychosomatique*.³⁵ (Perhaps influenced by Fonagy’s introduction of the concept, the term ‘mentalization’ would also enter the literature on social cognition in the 1990s to refer to the capacity to model the minds of others).³⁶

Related ideas were also circulating in British psychoanalytic circles, associated especially with the work of Bion, who had described distortions of social and self-perception stemming from problems in the representation of thoughts and feelings.³⁷ On Fonagy’s interpretation, Bion held that these distortions would be especially likely when an individual felt threatened, internally or externally. This helped Fonagy articulate the difference between mentalization and theory of mind, in a paper co-authored with the psychiatrist Anna Higgitt and published

³² Fonagy, P. (1989). ‘On Tolerating Mental States: Theory of Mind in Borderline Patients.’ *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 97.

³³ Fonagy, P. and Bateman, A. (2009). ‘A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice’, in Brownescombe Heller, M. and Pollet, S. (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176: ‘The first time we used the term “mentalization” was in 1989 (Fonagy 1989), influenced by the Ecole Psychosomatique de Paris, but we used the term as operationalized by developmental researchers investigating theory of mind (Leslie, 1987)’ (pp. 157–158). In fact, the term ‘mentalisation’ does not appear in Leslie: Leslie, A. M. (1987). ‘Pretense and Representation: The Origins of “theory of Mind”’. *Psychological Review*, 94(4): 412–426.

³⁴ Luquet, P. (1981). ‘Le Changement dans la Mentalisation’. *Revue Française de Psychoanalyse*, 45: 1023–1028.

³⁵ Marty, P. (1990). *Mentalisation et Psychosomatique: Les Empêcheurs de Penser en Rond*, Paris. In recent years, Fonagy and colleagues have returned to the topic of somatic disorders, bringing to the topic their developed theory of mentalization—distinct from the École Psychosomatique de Paris. See e.g. Ballespí, S., Vives, J., Alonso, N., Sharp, C., Ramírez, M. S., Fonagy, P., and Barrantes-Vidal, N. (2019). ‘To Know or Not to Know? Mentalization as Protection from Somatic Complaints’. *PloS One*, 14(5): e0215308; Luyten, P., De Meulemeester, C., and Fonagy, P. (2019). ‘Psychodynamic Therapy in Patients with Somatic Symptom Disorder’, in David Kealy and John S. Ogronczuk (eds), *Contemporary Psychodynamic Psychotherapy: Evolving Clinical Practice*, New York: Academic Press, pp. 191–206.

³⁶ Morton, J. and Frith, U. (1995). ‘Causal Modeling: A Structural Approach to Developmental Psychopathology’, in D. Cicchetti and D. J. Cohen (eds), *Manual of Developmental Psychopathology*, New York: Wiley, pp. 357–390: ‘The development of theory of mind (Premack and Woodruff, 1978) or “mentalising”—our ability to predict and explain the behaviour of other humans in terms of their mental states. Our ability to mentalise is revealed in our use and understanding of such words as believe, know, wish, desire, intend and pretend . . . Mentalising is primarily unconscious or implicit. It is a property of our cognitive apparatus that comes into action when triggered by particular stimuli, and it “makes sense” of other people’s and our own behaviour fully automatically’ (p. 363).

³⁷ The common element between Marty and Bion is likely Freud, S. ([1912] 2001). ‘Formulations on the Two Principles of Mental Functioning’, in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 12 (1911–1913): The Case of Schreber, Papers on Technique and Other Works, pp. 213–226, London: Vintage. To an extent, these concerns had been dealt with by Kleinian psychoanalysis under the rubric of ‘symbolization’. However, as Fonagy observed, the concept was insufficiently precise for fine work: Fonagy, P. (1991). ‘Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient’. *The International Journal of Psychoanalysis*, 72: 639–656: ‘The achievement of a representation of mental events, whether conscious or unconscious, is frequently referred to in the psychoanalytic literature in the context of the capacity for symbolization (Segal, 1957); (McDougall, 1974); (Edgcombe, 1984). The term symbolization is over-burdened with meanings, particularly in psychoanalysis. It is certainly not possible to restrict it to the notion of the secondary representation of mental states’ (p. 641).

in 1990, the same year that Fonagy and Higgitt married. Baron-Cohen and colleagues studying autism had conceptualized this condition as a pervasive deficit in theory of mind. By contrast, Fonagy and Higgitt characterized mentalization as the capacity to continue generating models of mental states in the face of perceived internal or external challenges or conflict:

Failures of mentalisation tend to be limited and specific to situations characterised by intense psychic conflict, particularly of an interpersonal kind. The large number of poets and writers (e.g. Baudelaire, Carroll, Nerval, Rimbaud and Verhaeren) whose biographies illustrate that their psychic functioning was close to or at the borderline level, yet whose capacity for abstract thinking in relation to the most subtle of human emotions could not be doubted, testify to this fact.³⁸

Whereas Baron-Cohen had conceptualized autism as a pervasive deficit in theory of mind, Fonagy and Higgitt characterized mentalization as, specifically, a capacity. It might not feature much of the time when not needed. Furthermore, like other capacities, we may sometimes experience lapses in our ability or desire to use it, perhaps especially under conditions of stress or anxiety. Nonetheless, Fonagy and Higgitt described the major potential value of having access to the capacity to identify and consider mental states for both psychological health and social relationships.

In a series of articles in the 1990s, Fonagy and Target elaborated implications of this perspective for clinical technique. They agreed with the observation of the borderline group at the Anna Freud Centre that psychoanalytic interpretations had the potential to backfire. They felt that they could pinpoint the reason for this: interpretations can only be helpful to patients who can think about the difference between i) their personal wishes and worries about their analyst, and ii) the analyst's here-and-now reality as an ally in reflection. Fonagy and Target also argued, however, against overreliance on pragmatic advice to patients about concrete actions. They instead encouraged the use of 'brief, accurate and simple statements of the analyst's perception of the patient's current mental state.'³⁹ There are resonances here of Fonagy's retrospective descriptions of his analysis with Anne Hurry. And, in the early 1990s, Fonagy echoed Hurry's advice that clinicians should adopt a tone of empathy with their patients. However, as the 1990s went on, Fonagy came to disagree on this point: 'I hate the word empathy because it carries the implication that you have to be nice.'⁴⁰ Fonagy felt that therapists could adopt a variety of tones in helping the patient gain perspective on mental states—these could be kind without being 'nice'. Hurry's own use of benevolent humour had suggested that what was important was the capacity of the therapist to spur the patient to mentalize, rather than empathy itself.

Fonagy and Target speculated that the active ingredient across most successful therapeutic modalities was actually support for improved mentalization. When patients in psychoanalytic psychotherapy could make use of interpretations, this would improve mentalization directly by allowing the patient to reflect on and integrate different perspectives on potentially distressing mental states. Cognitive Behavioural Therapy (CBT) was likewise anticipated to

³⁸ Fonagy, P. and Higgitt, A. (1990). 'A Developmental Perspective on Borderline Personality Disorder'. *Revue Internationale de Psychopathologie*, 1: 125–159, p. 140.

³⁹ Fonagy, P. and Target, M. (1995). 'Understanding the Violent Patient: The Use of the Body and the Role of the Father'. *The International Journal of Psychoanalysis*, 76: 487–501, p. 498.

⁴⁰ Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind* (pp. 77–98). Oxford: Butterworth Heinemann, p. 91.

contribute to mentalizing by helping patients elaborate and organize representations of their mental states, allowing them to be used appropriately even in the context of interpersonal or internal difficulties.⁴¹ For Fonagy and Target, enhanced mentalization would offer several benefits. It would support emotion regulation by interposing the capacity to reflect between the immediacy of felt experience and the chain of potential responses to these feelings. It would reduce anxiety by contributing to the orderliness of internal experience, and make this experience feel more manageable. And it would support the capacity of the patient to make use of social relationships to help them learn from experience, capitalize on potential opportunities for flourishing, and respond sensibly to adversities.

Reflective function

Interviewed in 2015, Fonagy was asked about the greatest regret of his career. He replied:

I spent the early years of my career a bit lost in the wilderness. It wasn't until two PhD students focused my attention on attachment and Bowlby that I found the topic in psychology that I was really genuinely interested in.⁴²

These students were Miriam and Howard Steele. The Steeles had been introduced to attachment theory by Larry Aber in New York.⁴³ Attachment theory in the late 1980s was undergoing an exciting revolution. Main and colleagues in Berkeley had introduced the Adult Attachment Interview in 1985, as a measure of a speaker's capacity to speak coherently about attachment-relevant memories. Main and colleagues found that some speakers were unable to maintain coherent discourse in the interview and would either close down the subject of attachment-relevant memories or become entangled in the memories and lose track of the interview questions. Another interruption of coherence could come from lapses in reasoning or discourse when speakers discussed experiences of bereavement or trauma, suggesting a disturbed state of mind regarding these events. The Adult Attachment Interview coding protocol is not concerned with assessing the speaker's childhood. There are sub-scales for inferred experience, such as the extent to which the child was rejected or neglected by their parents. However, the coding system primarily focuses on examining individual differences in *how* childhood attachment relationships are discussed by the adult speaker.

⁴¹ See Goodman, G., Midgley, N., and Schneider, C. (2016). 'Expert Clinicians' Prototypes of an Ideal Child Treatment in Psychodynamic and Cognitive-Behavioral Therapy: Is Mentalization seen as a Common Process Factor?' *Psychotherapy Research*, 26(5): 590–601. Recent work has found greater convergence in how CBT and psychoanalytic therapists respond to disengaged patients than how they respond to engaged patients. In terms of Fonagy and Hepworth's hypothesis that mentalizing is the active ingredient in both cases, it might be supposed that the modalities have different strategies for sustaining mentalizing—for instance, use of transference or homework—but both used similar approaches for bringing mentalizing online. e.g. 'they actively structured the sessions, asked many questions and tried to bring up material from previous sessions in an attempt to invite the young person to talk'. The researchers note that this strategy for responding to disengagement is mentioned in the treatment manuals of neither approach. Calderon, A., Schneider, C., Target, M., and Midgley, N. (2019). "Interaction Structures" between Depressed Adolescents and their Therapists in Short-Term Psychoanalytic Psychotherapy and Cognitive Behavioural Therapy'. *Clinical Child Psychology and Psychiatry*, 24(3): 446–461, p. 457.

⁴² Fonagy, P. (2015). 'I Would Like to Abolish Silo Working' *The Psychologist*, 28: 948. Accessed at: <https://thepsychologist.bps.org.uk/i-would-abolish-silo-working>.

⁴³ Steele, H. and Steele, M. (2005). 'Understanding and Resolving Emotional Conflict: The London Parent-Child Project', in K. E. Grossmann, K. Grossmann, and E. Waters (eds), *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*, New York: Guilford Press, pp. 137–164.

Main and colleagues found that speakers lower in coherence on the Adult Attachment Interview were more likely to have relationships with their child classified as insecure in the Ainsworth Strange Situation.⁴⁴ The Strange Situation is an observational procedure in which repeated separations and reunions of an infant from their familiar caregiver are inferred to activate the attachment system, prompting the infant to want to seek the availability of their caregiver. Individual differences in response to the activation of the attachment system are observed. 'Secure attachment' is coded when infants communicate their distress at the separation but can be comforted by reunion and able to then return to play. Ainsworth termed this pattern the use of the caregiver as a 'secure base' and 'safe haven', because it seemed to reflect trust in the caregiver's availability. Some infants did not display their distress to the caregiver; Ainsworth found that these infants experienced rebuff of their proximity-seeking behaviours at home, and she termed the dyads 'avoidantly attached'. Some infants showed frustration and were unable to accept comfort from the caregiver; Ainsworth found that these infants had caregivers who were not adept at responding accurately to their infants' cues.⁴⁵ Ainsworth labelled these dyads 'ambivalent/resistantly attached'. A further classification for infant behaviour in the Strange Situation, introduced by Main and colleagues, was 'disorganized' attachment. Main and colleagues characterized disorganized attachment as representing the most concerning form of insecure behaviour.

It should be pointed out that the concept of disorganized attachment was introduced by Main and colleagues in quite a misleading way. The term 'disorganization' in ordinary language means something thrown into chaos. By contrast, they used the term in a technical sense, but without explaining this to the reader.⁴⁶ Disorganized attachment was coded on the basis of infant displays of i) conflicted, ii) confused, or iii) apprehensive behaviour towards their caregiver, where these suggest conflict in expectations about the availability of the caregiver as a safe haven. When a child's behaviour in the Strange Situation actually appears random, this is coded as 'Cannot Classify', rather than disorganization. The assignment of disorganized attachment actually requires the coder to observe determinate 'patterning within the disturbance of coherence', rather than unpredictability.⁴⁷ Another source of confusion was that Main and Hesse proposed that disorganized attachment is caused by fear in relation to the caregiver. This suggests a unitary process, triggered by hostile caregiving. Certainly disorganized attachment was, soon after its introduction, identified as having increased prevalence in maltreatment samples.⁴⁸ However, Main and Hesse found that the

⁴⁴ Main, M., Kaplan, N., and Cassidy, J. (1985). 'Security in Infancy, Childhood, and Adulthood: A Move to the Level of Representation'. *Monographs of the Society for Research in Child Development*, 50(1–2): 66–104.

⁴⁵ Ainsworth, M. D. S. (1983). 'Patterns of Infant–Mother Attachment as Related to Maternal Care: Their Early History and their Contribution to Continuity', in D. Magnusson and V. L. Allen (eds), *Human Development: An Interactional Perspective*, New York: Academic Press, pp. 35–55.

⁴⁶ The challenge for psychological theory was evoked well by Bion in his Brazil lectures: 'If we invent words, nobody will understand what those words mean. If we do not, their sensuous history is evoked. As with a "dead" metaphor carelessly used, its ghost begins to walk.' Bion, W. R. ([1973] 1990). *Brazilian Lectures*, London: Karnac Books, p. 34. And, in any case, 'attempts to contrive a term devoid of a distorting penumbra of associations is often defeated by the speed with which such a meaningless term collects a meaning.' Bion, W. R. (1962) *Elements of Psychoanalysis*, London: Karnac Books, p. 88. As a thought experiment, if Main and Solomon had set out their observations using a technical vocabulary rather than the evocative language of 'disorganisation' and 'fear', it seems likely that it would have had only a fraction of the take-up outside attachment research. At the same time, a price was paid in simplifying the concepts beyond all recognition. Duschinsky, R. (2018). 'Disorganization, Fear and Attachment: Working towards Clarification'. *Infant Mental Health Journal*, 39(1): 17–29.

⁴⁷ Reijman, S., Foster, S., and Duschinsky, R. (2018). 'The Infant Disorganised Attachment Classification: "Patterning within the Disturbance of Coherence"', *Social Science & Medicine*, 200: 52–58.

⁴⁸ Carlson, V., Cicchetti, D., Barnett, D., and Braunwald, K. (1989). 'Finding Order in Disorganization: Lessons from Research on Maltreated Infants' Attachments to their Caregivers', in D. Cicchetti and V. Carlson (eds), *Child Maltreatment: Theory and Research on the Causes and Consequences of Child Abuse and Neglect*, New York: Cambridge University Press, pp. 494–528.

caregivers from dyads classified as showing disorganized attachment were especially likely to be unresolved for bereavement or trauma on the Adult Attachment Interview, among other antecedents.⁴⁹ Such findings suggested that the Adult Attachment Interview could be a powerful tool for exploring factors that disrupt the development of secure child–caregiver attachment.

Miriam and Howard Steele brought the first article on the Adult Attachment Interview with them when they moved to London and shared this with Fonagy. Fonagy was especially excited by the fact that ‘both attachment theory and modern psychoanalysis have as their fundamental epistemic aim the description of the internal mechanisms responsible for the discrepancy between actual and psychic reality.’⁵⁰ The Steeles and Fonagy attended training in the Adult Attachment Interview held in 1987 at the Tavistock Clinic, taught by Mary Main and Erik Hesse at the invitation of John Bowlby.⁵¹ Between 1987 and 1989, the Steeles conducted the Adult Attachment Interview with 100 expectant mothers and fathers in the third trimester. They also collected 96 infant–mother Strange Situation observations when the children were 12 months old, and 90 infant–father Strange Situation observations at 18 months. This was the basis for the University College London Parent–Child Project, a fascinating longitudinal study that would eventually follow the sample over the subsequent 17 years.

The first findings from the UCL Parent–Child Project were published by Fonagy, Steele, Moran, Steele, and Higgitt in 1991. The researchers reported powerful intergenerational continuities for mothers: around three-quarters of children whose mothers showed low coherence on the Adult Attachment Interview, conducted prenatally, displayed insecure patterns of attachment in the Strange Situation procedure. The same was true for around half of fathers. Eighty per cent of children whose mothers or fathers had high coherence in the Adult Attachment Interview displayed a secure pattern of attachment in the Strange Situation. The Strange Situation classifications were not predicted by self-report of the quality of the marital relationship or by attitudes towards caregiving by either parent.⁵² The coding system for the Adult Attachment Interview had a sub-scale for ‘meta-cognitive monitoring’, the extent to which a speaker seemed able to reflect on their own thoughts, feelings, and discourse. However, the scale was (and, in fact, still is) unfinished. The Steeles and Fonagy, examining the transcripts of coherent speakers with a clinical eye, ‘noticed that a defining feature of their narratives was the way they relied on language as a tool for giving meaning to experience, including

⁴⁹ Main, M. and Hesse, E. (1990). ‘Parents’ Unresolved Traumatic Experiences are Related to Infant Disorganized Attachment Status: Is Frightened and/or Frightening Parental Behavior The Linking Mechanism?’, in M. T. Greenberg, D. Cicchetti, and E. M. Cummings (eds), *Attachment in the 638 Child Development Preschool Years*, Chicago: University of Chicago Press, pp. 161–182. On the antecedents of disorganized attachment, see also Cyr, C., Euser, E. M., Bakermans-Kranenburg, M. J., and Van IJzendoorn, M. H. (2010). ‘Attachment Security and Disorganization in Maltreating and High-Risk Families: A Series of Meta-Analyses’. *Development and Psychopathology*, 22(1): 87–108; Granqvist, P., Sroufe, L. A., Dozier, M., Hesse, E., Steele, M., van IJzendoorn, M., ... and Duschinsky, R. (2017). ‘Disorganized Attachment in Infancy: A Review of the Phenomenon and its Implications for Clinicians and Policy-Makers’. *Attachment & Human Development*, 19(6): 534–558.

⁵⁰ Fonagy, P. (1999). ‘Points of Contact and Divergence between Psychoanalytic and Attachment Theories: Is Psychoanalytic Theory Truly Different?’ *Psychoanalytic Inquiry*, 19(4): 448–480, p. 452.

⁵¹ Steele, H. and Steele, M. (2005). ‘Understanding and Resolving Emotional Conflict: The London Parent–Child Project’, in K. E. Grossmann, K. Grossmann and E. Waters (eds), *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*, New York: Guilford Press, pp. 137–164.

⁵² Fonagy, P., Steele, M., Moran, G. S., Steele, H., and Higgitt, A. C. (1991). ‘Measuring the Ghost in the Nursery: A Summary of the Main Findings of the Anna Freud Centre/University College London Parent–Child Study’. *Bulletin of the Anna Freud Centre*, 14: 115–131.

the attribution of mental states (beliefs and desires).⁵³ These qualities seemed relevant across the Adult Attachment Interview, but especially in response to two questions that demanded examination of the motivations and influences guiding the behaviour of the speaker and their attachment figures: ‘When you think about your childhood experiences, do you think they have an influence on who you are today?’ and ‘When you think about your parents’ behaviour toward you when you were a child, why do you think they behaved the way they did?’⁵⁴

The Steeles and Fonagy initially ‘referred to this phenomenon as evidence of the “internal observer” as we thought this term captured the sense in which adult speakers could observe how family life was when they were children, and distinguish this from the understanding they gained through later relationships across development.’⁵⁵ However, the ‘internal observer’ concept was amended to the ‘reflective self’.⁵⁶ This was a term from the psychoanalyst Christopher Bollas, used to encompass the fact that the attribution and reconsideration of mental states seemed to be more than perspective-taking on the past, and could include appraisal of different perspectives in the present or imagined future.⁵⁷ The term was then extended again to ‘reflective function’, to encompass the capacity of the speaker to interpret the experiences of attachment figures with reference to mental states.

The Adult Attachment Interview had been modelled on a clinical interview, integrating two components. The earlier component was a psychoanalytic approach to clinical interviewing from the 1970s, in which psychological defences or confusions were deciphered by the clinician. The later component, under the influence of the American Psychiatric Association’s DSM-III, was use of clinical interviews from the 1980s to identify specifiable markers for categories of mental pathology within the particularities of what and how an individual reports their experiences.⁵⁸ Reflective function can likewise be seen as a further integration of the Adult Attachment Interview with developments within psychoanalysis that turned attention to obstacles to thinking and understanding (see Chapter 1). Indeed, though not situated as such by Fonagy and colleagues at the time, it is very likely that an influence on the concept of ‘reflective function’ was Bion’s earlier concept of ‘alpha function’, to mean the process by which determinate thoughts and feelings are made out of experiences.⁵⁹

⁵³ Steele, H. and Steele, M. (2005). ‘Understanding and Resolving Emotional Conflict: The London Parent–Child Project’, in K. E. Grossmann, K. Grossmann, and E. Waters (eds), *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*, New York: Guilford Press, pp. 137–164, p. 157.

⁵⁴ Steele, H. and Steele, M. (2008). ‘On the Origins of Reflective Functioning’, in F. Busch (ed.), *Mentalization: Theoretical Considerations, Re-Search Findings, and Clinical Implications*, New York: Taylor and Francis, pp. 133–158.

⁵⁵ Steele, H. and Steele, M. (2005). ‘Understanding and Resolving Emotional Conflict: The London Parent–Child Project’, in K. E. Grossmann, K. Grossmann, and E. Waters (eds), *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*, New York: Guilford Press, pp. 137–164, p. 157.

⁵⁶ As in the subtitle of Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). ‘The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment’. *Infant Mental Health Journal*, 12(3): 201–218.

⁵⁷ This usage was influenced by Bollas, C. (1990). ‘The Origins of the Therapeutic Alliance’. Paper presented at the English Speaking Conference of the British Psychoanalytical Society, October, London. Later published as Bollas, C. (1998). ‘Origins of the Therapeutic Alliance’. *Scandinavian Psychoanalytic Review*, 21(1): 24–36.

⁵⁸ Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press.

⁵⁹ Fonagy and Target would acknowledge alignment between the idea of reflective function and that of alpha function, though not any relation of influence e.g. Fonagy, P., Target, M., Steele, H., and Steele, M. (1998). *Reflective-Functioning Manual*, Version 5, London: University College London/Anna Freud Centre, p. 4; Fonagy, P. and Target, M. (2003). ‘Evolution of the Interpersonal Interpretive Function: Clues for Effective Preventive Intervention in Early Childhood’, in S. W. Coates, J. L. Rosenthal, and D. S. Schechter (eds), *September 11: Trauma and Human Bonds*, Hillsdale, NJ: Analytic Press, pp. 99–113, p. 102. The developmental model of reflective functioning likewise has marked similarities to the model proposed by Bion. Though since Fonagy and Target do not acknowledge influence, it is difficult to know whether this can be inferred. See e.g. Bion, W.R. ([1973] 1990). *Brazilian Lectures*, London: Karnac Books: ‘If the baby has a mother who is able to tolerate being frightened that the child is

Frequently, reflective function and mentalization were treated as synonyms by Fonagy and colleagues in the 1990s and 2000s. However, a few distinctions can be identified. Whereas the term 'mentalization' was used by Fonagy and colleagues to refer to the capacity of an individual to think about mental states, the term 'reflective function' appears to have been preferred, especially by the Steeles, to describe the capacity to think about mental states in the context of attachment relationships specifically.⁶⁰ The Steeles and Fonagy developed a manual for coding reflective functioning and applied it to the transcripts. Though this was the development sample for the measure, and so strong findings might be expected, the results were still surprisingly clear. Reflective function in the Adult Attachment Interview not only predicted security in the Strange Situation better than coherence for mothers and fathers, but coherence had no association with security when reflective functioning was statistically controlled.⁶¹ The Steeles and Fonagy also checked how the reflective functioning measure compared with the inferred experience scales: again, they found that reflective functioning was a better predictor than the inferred experience scales, which had no association with the Strange Situation once reflective functioning was controlled. The researchers concluded:

The coherence of the parents' perception of their past derives from their unhindered capacity to observe their own mental functioning, to have a plausible view of themselves and their objects as human beings, thinking, feeling, wishing, believing, wanting, and desiring. We assume, then, that coherence may be a measure of reflectiveness, and it is the latter attribute of the caregivers that has direct implications for their relationship to the infant . . . In our view, a child may be said to be secure in relation to a caregiver to the extent that, on the basis of his or her experience, he or she can make an assumption that his or her mental state will be appropriately reflected on and responded to accurately. The child's confidence in this assumption will enable her or him to be more assured about the safety of the world of ideas and desires; she or he will feel secure in relation to her or his mental world.⁶²

The unpublished manual conveys in more detail how the construct of reflective function was operationalized. It was by no means a quantitative count of occasions when mental state terms were used by speakers:

for a statement to be reflective, it has to imply an activity of reflection, rather than simply describing a mental state. Thus, the statement 'I was angry' would not be considered

dying, or that it is mad, or stupid, then the baby seems to feel better for having such a mother. If she cannot tolerate it, neither can the baby who consequently seems unable to grow mentally. If it is going to grow it has to do so in a peculiar way; this, in turn, has the effect of making it grow a certain shape. Later, someone will say . . . "borderline case" (p. 132).

⁶⁰ Sharp, C. and Fonagy, P. (2008). 'The Parent's Capacity to Treat the Child as a Psychological Agent: Constructs, Measures and Implications for Developmental Psychopathology'. *Social Development*, 17(3): 737–754: 'What exactly does mentalizing mean within this context? . . . When this capacity is operationalized within the context of attachment relationships, Fonagy refers to it as "reflective functioning" (RF) . . . The term is used to describe the parent's capacity to reflect upon his/her own or the child's internal mental experience within the context of attachment' (p. 740).

⁶¹ Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). 'The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment.' *Infant Mental Health Journal*, 12(3): 201–218.

⁶² *Ibid.* 215.

reflective, whereas the statement 'At the beginning I felt really angry and then it changed to sadness' does suggest that the feelings have been considered.⁶³

Compared with the general construct of mentalization, the operationalization of reflective function is notable for a few subtle differences, perhaps shaped by the priorities of the Adult Attachment Interview. The coding of reflective function appeared to give special weight to the capacity of the speaker to identify change in their perspective over time. And, more than the idea of mentalization, reflective function was concerned with understanding behaviour and mental states as having *bidirectional* lines of influence, not solely the interpretation of mental states as the cause of behaviour.⁶⁴

The manual advised that individual statements should be given a score on a 1–9 scale, as should the transcript as a whole. This required the coder to make a synthesis of the extent, intensity, and quality of the varied forms of reflective functioning evident across the transcript. Fonagy and colleagues reported that 'when reflective-functioning is absent or ineffectual, quite distinct, readily classifiable, patterns of responses tend to emerge.'⁶⁵ One pattern was rejection of reflective functioning, with the prompts in the interview to think about childhood experiences and the intentions of attachment figures responded to in a hostile way by the speaker. Another pattern of particular note was 'hyperactive reflective function', in which there is extensive discussion of mental states, but these are poorly integrated and do not have much bearing on actually answering the questions posed by the interviewer: 'Mentalization spins like a car wheel which has lost contact with ground. Because so many of the connections between what is thought about thoughts or feelings and the original thoughts or feelings themselves have been lost, metarepresentational capacity works overtime but without real impact.'⁶⁶ This kind of response would later be termed 'hypermentalizing' (see Chapter 5).

A further pattern identified by Fonagy and colleagues was 'disavowal of reflective function', the most common form of which were explanations of caregivers' behaviour in 'sociological' rather than psychological terms, such as 'people in their social class never expected to be affectionate with their children'. Fonagy had earlier excluded Dunn's attention to social convention as part of mentalization. Taking this logic further, now the operationalization of reflective function actively treated sociological explanations as rather at the expense of reflection of the mental states of determinate figures. Mentalization in general, and reflective function in particular, were conceived of as 'psychological processes underpinning the

⁶³ Fonagy, P., Target, M., Steele, H., and Steele, M. (1998). *Reflective-Functioning Manual*, Version 5, London: University College London/Anna Freud Centre.

⁶⁴ This point seems to be explicitly affirmed in Murphy, A., Steele, M., and Steele, H. (2013). 'From Out of Sight, Out of Mind to In Sight and In Mind: Enhancing Reflective Capacities in a Group Attachment-Based Intervention', in Bettmann Joanna Ellen and Friedman Donna Demetri (eds), *Attachment-Based Clinical Work with Children and Adolescents*, NY: Springer, pp. 237–257: 'We define reflective functioning as (1) an awareness of the nature of mental states in the self and others; (2) the mutual influences at work between mental states and behaviour; (3) the necessity of a developmental perspective; and (4) the need to be sensitive to the current context' (pp. 238–239). Discrepancy between the characterizations of reflective function and mentalization, despite claims by Fonagy and colleagues that they relate to the same construct, have also been noted by Eagle, M. (2013). *Attachment and Psychoanalysis*, New York: Guilford Press, p. 184. For Eagle, mentalization requires treating thoughts and feelings as intentions; reflective function addresses thoughts and feelings without the assumption of intentionality. Eagle also sees mentalization as implicit compared with reflective function as an explicit process. In our view, both contrasts are important, and there are some uses of the concepts that fit this contrast. However, at other times, reflective function is treated as implicit and assuming intentionality, and mentalization is treated as explicit and concerned with thoughts and feelings in general. Eagle's characterization therefore does not hold, though he is certainly correct to identify these two axes of underlying ambiguity.

⁶⁵ Fonagy, P., Target, M., Steele, H., and Steele, M. (1998). *Reflective-Functioning Manual*, Version 5, London: University College London/Anna Freud Centre.

⁶⁶ *Ibid.*

view of oneself and others as motivated by mental states.’⁶⁷ Attention to processes above the level of mental states, such as sociological forces were therefore characterized as the absence of mentalizing (Chapter 9). Embodied processes below the level of mental states, such as endocrinology (e.g. appeal to the causal role of adrenalin or cortisol; consideration of post-partum hormone changes), are also ambiguous in this regard.⁶⁸ The formulation offered by Fonagy and colleagues would seem to implicitly exclude them as legitimate objects of mentalization.⁶⁹

The link between caregiver reflective function and infant attachment security in the UCL Parent–Child Project suggested to Fonagy and colleagues four revisions or amendments to important concepts in attachment theory in the late 1990s:

- Sensitive care.
- The cycle of abuse.
- Internal working model.
- Security.

First, Fonagy and colleagues proposed a new perspective on sensitive care. Ainsworth had proposed that the basis for secure infant attachment lay in caregiver ‘sensitivity’, a technical term she used for the capacity to ‘perceive and to interpret accurately the signals and communications implicit in her infant’s behavior, and given this understanding, to respond to them appropriately and promptly.’⁷⁰ Though incredibly perspicacious, there are nonetheless several problems with this account. It assumes that infants offer a single signal at a time or that signals are ultimately compatible; that it is in the parent’s capacity to solve the problem; that there are no needs without, prior to, or contrary to signals; that needs are not incompatible; and it neglects the importance of distraction.⁷¹ Perhaps most importantly, not all signals pertain to attachment needs. An infant may cry because they are bored or confused, for example. So even exquisite sensitivity will only ever be at best a moderate predictor of

⁶⁷ Fonagy, P. (1996). ‘The Significance of the Development of Metacognitive Control over Mental Representations in Parenting and Infant Development’. *Journal of Clinical Psychoanalysis*, 5(1): 67–86, p. 74.

⁶⁸ Cf. Cassidy, J., Jones, J. D., and Shaver, P. R. (2013). ‘Contributions of Attachment Theory and Research: A Framework for Future Research, Translation, and Policy’. *Development and Psychopathology*, 25(4.2): 1415–1434: ‘Consideration of linkages between representational and nonrepresentational processes must include the possibility that causality flows in both directions: physiological stress responses can presumably prompt a person to engage in higher level cognitive processes to understand, justify, or eliminate the stressor’ (p. 1419).

⁶⁹ The exception would be where conventions are established to infer mental states from physiology or endocrinology. E.g. McHugh, B., Dawson, N., Scrafton, A., and Asen, E. (2010). ‘Hearts on Their Sleeves’: The Use of Systemic Biofeedback in School Settings’. *Journal of Family Therapy*, 32(1): 58–72; Fotopoulou, A. and Tsakiris, M. (2017). ‘Mentalizing Homeostasis: The Social Origins of Interoceptive Inference’. *Neuropsychanalysis*, 19(1): 3–28. The distinction between the physical and the mental in the case of gender has been described by Fonagy as ‘wrong’ and ‘inappropriate’. However, he has not yet been clear about whether attention to embodied processes below the level of mental states would constitute mentalizing. See Fonagy’s remarks in Steidinger, S. (2018) ‘Trans-Actions: An Exploration of Gender Dysphoria’. Accessed at: <https://vimeo.com/285555219>.

⁷⁰ Ainsworth, M. (1969). ‘Scale 1: Sensitivity vs. Insensitivity to the Baby’s Signals’. Accessed at: http://www.psychology.sunysb.edu/attachment/measures/content/ainsworth_scales.html. Some but not all aspects of the sensitivity scale and construct were described in Stayton, D. J., Hogan, R., and Ainsworth, M. (1971). ‘Infant Obedience and Maternal Behavior: The Origins of Socialization Reconsidered’. *Child Development*, 42(4): 1057–1069, pp. 1060–1061.

⁷¹ Some of these concerns are discussed variously in Kondo-Ikemura, K. (2001). ‘Insufficient Evidence’. *American Psychologist*, 56(10): 825; Keller, H. and Otto, H. (2009). ‘The Cultural Socialization of Emotion Regulation During Infancy’. *Journal of Cross-Cultural Psychology*, 40(6): 996–1011; Shai, D. and Belsky, J. (2017). ‘Parental Embodied Mentalizing: How the Nonverbal Dance between Parents and Infants Predicts Children’s Socio-Emotional Functioning’. *Attachment & Human Development*, 19(2): 191–219.

attachment security.⁷² For their part, Fonagy and colleagues argued that caregiver sensitivity was largely a by-product of reflective function, and that it was the latter that was largely the basis for infant secure attachment. What was captured by Ainsworth's sensitivity scale, Fonagy and colleagues proposed, was the extent to which the caregiver was attentive to the mental states of the child, could interpret these accurately, and could respond to them promptly and in an undistorted way.⁷³ Later research has indeed supported the claim that some of the association between sensitivity and infant attachment can be accounted for in terms of parental reflective function.⁷⁴ However, it would have been surprising if all of the association had been accounted for, because both sensitive care and attachment security can readily be seen in non-human animals with far less role for representation in parent-child interactions.⁷⁵ There may be different paths to sensitivity through automatic and controlled forms of caregiver mentalizing (see Chapter 4)—but this was not a distinction Fonagy and colleagues had available in the late 1990s, nor one they have subsequently considered with respect to sensitivity.

A second alteration to attachment theory was regarding the 'cycle of abuse'. This was a concept introduced by Egeland, Jacobvitz, and Sroufe in 1984, to describe findings from the Minnesota longitudinal study that about a third of caregivers who reported that they had been abused went on to abuse their own child: a much higher proportion than matched controls, even if overall a minority of cases.⁷⁶ Egeland and colleagues reported that the abused mothers who did not go on to abuse their child were distinguished by one or more of three experiences. First, some had emotional support from a non-abusive adult during their childhood. Second, some had participated in therapy. Third, some had a stable and emotionally supportive relationship with their adult partner. Fonagy and colleagues argued that what these experiences had in common was that they could be expected to increase a caregiver's reflective functioning, offering possibilities for interrupting responses that might otherwise predispose abusive or neglecting behaviour towards their child.⁷⁷

⁷² Woodhouse, S. S., Scott, J. R., Hepworth, A. D., and Cassidy, J. (2020). 'Secure Base Provision: A New Approach to Examining Links Between Maternal Caregiving and Infant Attachment'. *Child Development*, 91(1): e249–e265.

⁷³ Fonagy, P., Steele, H., Steele, M. and Holder, J. (1997). 'Attachment and Theory of Mind: Overlapping Constructs?' *Association for Child Psychology and Psychiatry Occasional Papers*, 14: 31–40; for Ainsworth, 'the caregiver's perception of the child as an intentional being lies at the root of sensitive caregiving' (p. 36).

⁷⁴ Evaluating the criticisms of Ainsworth by Fonagy and colleagues, Zeegers and colleagues conducted a meta-analysis to investigate the relative contributions of sensitivity and mentalization/reflective function to infant attachment classifications. They found that together the two predictors accounted for 12% of variance in attachment classifications. After controlling for sensitivity, the relationship between parental mentalization and infant-caregiver security was $r = .24$. And, after controlling for mentalization, the relationship between parental sensitivity and infant-caregiver security was $r = .19$. Sensitivity also partially mediated the association between mentalization and infant-caregiver security ($r = .07$). Such findings suggest that Fonagy and others were right to argue for the importance of mentalization/reflective function, but that sensitivity is not reducible to reflective function. Zeegers, M. A., Colonnese, C., Stams, G. J. J., and Meins, E. (2017). 'Mind Matters: A Meta-Analysis on Parental Mentalization and Sensitivity as Predictors of Infant-Parent Attachment'. *Psychological Bulletin*, 143(12): 1245–1272.

⁷⁵ Bowlby, K. (1991). 'Ethological Light on Psychoanalytical Problems', in P. Bateson (ed.), *The Development and Integration of Behaviour: Essays in Honour of Robert Hinde*, Cambridge: Cambridge University Press, pp. 301–313; Stevenson-Hinde, J. (2005). 'The Interplay between Attachment, Temperament and Maternal Style: A Madingly Perspective', in K. E. Grossmann, K. Grossmann and E. Waters (eds.), *Attachment from Infancy to Adulthood: The Major Longitudinal Studies*, New York: Guilford Press, pp. 198–222.

⁷⁶ Egeland, B., Jacobvitz, D., and Sroufe, L. (1988). 'Breaking the Cycle of Abuse'. *Child Development*, 59(4): 1080–1088.

⁷⁷ Offering some support for this conclusion, researchers would find a large correlation between caregivers' reflective functioning in the Adult Attachment Interview and the absence of frightening, dissociative, anomalous, withdrawing behaviours or severe disruptions to parent-child communication during filmed observation. Fonagy, P. and Target, M. (2005). 'Bridging the Transmission Gap: An End to an Important Mystery of Attachment Research?' *Attachment & Human Development*, 7(3): 333–343.

A third change to attachment theory in the late 1990s was regarding the concept of ‘internal working model’. This concept had been introduced by Bowlby to characterize the expectations held by a child about the availability of their caregiver. Fonagy held that the concept of ‘internal working model’ was too vague, given that it encompassed memories of previous interactions with attachment figures, expectations about interactions with attachment figures, ramifications of these memories and expectations for a child’s sense of self, and ramifications of these memories and expectations for the child’s sense of their caregiver.⁷⁸ When attachment researchers proposed that internal working models were the mechanism through which early childhood experiences contributed to later behaviour, Fonagy regarded this as a badly underspecified claim. He and his father had been discussing the ideas of the philosopher Hegel, who had emphasized the integral role of the other in the self’s recognition of itself.⁷⁹ And from György Gergely, Fonagy took the idea that self-recognition would be facilitated when the caregiver’s behaviour provided cues to the child that their mental states and intentions were being acknowledged. Gergely argued that an especially potent form of recognition a child could receive is ‘marked mirroring’. This is when a caregiver provides enough of a mirror to the child’s behaviour to show that the intention and/or mental state implicated in the behaviour has been acknowledged, alongside some slight transformation or elaboration to also communicate a sense of acknowledgement. So, for instance, a young child’s cry may be met initially with a sad mouth and tonal echo from the parent, signalling that the emotion has been received, but this response is then transformed by a half-smile and a rising-pitch intonation, as if to convey that the sadness is localized, resolvable, and will be okay soon.⁸⁰ Marked mirroring utilizes the interplay of primary and secondary meanings highlighted by Ivan Fónagy as integral to human communication (see Chapter 1), offering the primary meaning of acceptance and the secondary meaning of containment.

⁷⁸ Fonagy, P. (2001). ‘The Human Genome and the Representational World: The Role of Early Mother–Infant Interaction in Creating an Interpersonal Interpretive Mechanism’. *Bulletin of the Menninger Clinic*, 65(3): 427–448, p. 436.

⁷⁹ Fónagy, I. and Fonagy, P. (1995). ‘Communication with Pretend Actions in Language, Literature and Psychoanalysis’. *Psychoanalysis and Contemporary Thought*, 18(3): 363–418. In the later formulation of Bateman, A. W. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorders: Mentalization-Based Treatment*, Oxford: Oxford University Press: ‘We have to assume a dialectic model of self-development (Hegel 1807) where the child’s capacity to create a coherent image of mind is critically dependent on an experience of being clearly perceived as a mind by the attachment figure’ (p. 64).

⁸⁰ Gergely, G. and Watson, J. (1996). ‘The Social Biofeedback Model of Parental Affect Mirroring. The Development of Emotional Self-Awareness and Self-Control in Infancy’. *The International Journal of Psychoanalysis*, 77: 1–31. A difference between Gergely and Fonagy in this regard is that the latter quite frequently translated the idea of marked mirroring into the Kleinian vocabulary of the caregiver offering ‘containment’ for the child’s affect. Kleinians, especially Bion, had certainly offered aligned ideas over the previous decades, though usually phrased in the language of projective identification. For instance, Bion wrote that ‘An understanding mother is able to experience the feeling of dread, that this baby was striving to deal with by projective identification, and yet retain a balanced outlook. This patient had had to deal with a mother who could not tolerate experiencing such feelings and reacted either by denying them ingress, or alternatively by becoming a prey to the anxiety which resulted from introjection of the infant’s feelings ... a well-balanced mother can accept these and respond therapeutically: that is to say in a manner that makes the infant feel it is receiving its frightened personality back again but in a form that it can tolerate’ (Bion, W. R. (1962). ‘The Psycho-Analytic Study of Thinking’. *The International Journal of Psychoanalysis*, 43: 306–310, p. 308. One subsequent criticism that Fonagy and colleagues have raised regarding the concept of marked mirroring is that it underplays the irreducibly kinaesthetic aspects of the process: Shai, D. and Fonagy, P. (2014). ‘Beyond Words: Parental Embodied Mentalizing and the Parent Infant Dance’, in M. Mikulincer and P. R. Shaver (eds), *Mechanisms of Social Connections: From Brain to Group*, Washington, DC: American Psychological Association, pp. 185–203. More than marked mirroring, the revised account in Shai and Fonagy emphasizes repair of interactive missteps as the paradigmatic form in which parental reflective function contributes to child attachment security, felt coherence, and socio-emotional development. See also Shai, D. and Meins, E. (2018). ‘Parental Embodied Mentalizing and its Relation to Mind-Mindedness, Sensitivity, and Attachment Security’. *Infancy*, 23(6): 857–872.

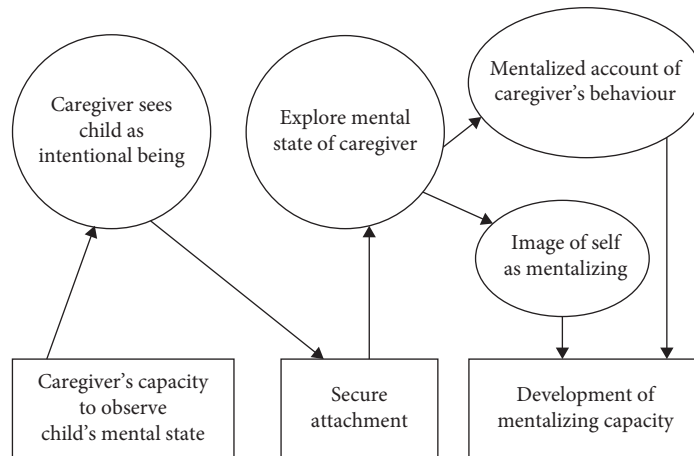


Figure 3.1 Developmental model ('paraphrased from Hegel').

Source: Reproduced from Peter Fonagy, 'Prevention, the Appropriate Target of Infant Psychotherapy', *Infant Mental Health Journal*, 19(2): 124–150, Figure 4, DOI: [https://doi.org/10.1002/\(SICI\)1097-0355\(199822\)19:2<124::AID-IMHJ4>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0355(199822)19:2<124::AID-IMHJ4>3.0.CO;2-O) Copyright © 1998 Michigan Association for Infant Mental Health.

Drawing on these ideas, Fonagy and colleagues proposed that patterns of attachment should be regarded as a correlate of the child's early mentalizing capacities, achieved through recognizing themselves in the minds of their caregiver. He felt that attachment researchers had reified the infant classifications and failed to truly develop proposals about 'the mechanisms or psychic processes that may underlie such behavioral clusters'.⁸¹ He speculated that secure attachment reflected a child's experiences of being treated by their caregiver as having intentions and/or mental states. Avoidant and ambivalent attachment reflected obstacles in the caregiver's recognition of their need for a secure base or a safe haven, and in turn could be anticipated to contribute to deficits in reflective function: "The avoidant child shuns the mental state of the other, while the resistant child focuses on its own state of distress to the exclusion of intersubjectivity".⁸²

Yet despite such qualitative differences, Fonagy and colleagues argued that the underlying mechanisms were likely best characterized in terms of a dimension. Ultimately, for Fonagy and colleagues, the importance of individual differences in attachment was that they contributed to or hindered the child's capacity for mentalization. It was this, not attachment itself nor the representations of caregivers that stem from them, that they felt contributed to a child's later socio-emotional development. Fonagy and Target argued that 'the influence of attachment security on later development has nothing to do with representations of early relationships, and a futile search for this link has distracted attachment researchers'.⁸³ The transition in the stance of Fonagy and colleagues can be seen in comparing the developmental model in a 1998 diagram with one 10 years later. Attachment security still figures in the second diagram, but it is not treated as a necessary path to child mentalising, nor necessary for a child's emotion regulation or later mental health.

⁸¹ Fonagy, P. (1999). 'Points of Contact and Divergence between Psychoanalytic and Attachment Theories: Is Psychoanalytic Theory Truly Different?' *Psychoanalytic Inquiry*, 19(4): 448–480, p. 469.

⁸² Ibid. 460.

⁸³ Fonagy, P. and Target, M. (2002). 'Early Intervention and the Development of Self-Regulation'. *Psychoanalytic Inquiry*, 22(3): 307–335, p. 328.

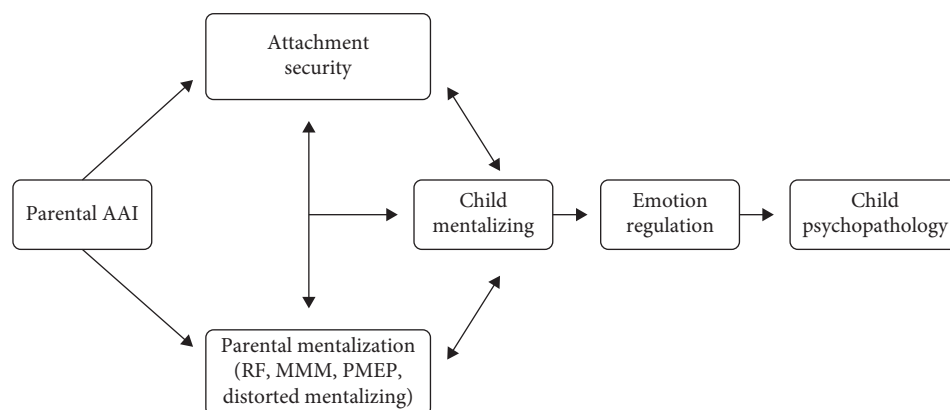


Figure 3.2 Developmental model.

Source: Reproduced from Carla Sharp and Peter Fonagy, 'The Parent's Capacity to Treat the Child as a Psychological Agent: Constructs, Measures and Implications for Developmental Psychopathology', *Social Development*, 17(3): 737–754, DOI: <https://doi.org/10.1111/j.1467-9507.2007.00457.x> Copyright © Blackwell Publishing Ltd. 2008.

Finally, a fourth revision to attachment theory proposed by Fonagy and colleagues addressed the concept of security and its clinical implications. Bowlby and several other attachment theorists had argued that the therapist offered a 'secure base' to the patient, from which they could venture to explore difficult thoughts and feelings. This process activated the patient's expectations about attachment relationships, and at the same time provided an opportunity to reflect on them. In this way, the therapy could contribute to revision in internal working models and a potential increase in the security of these models. Fonagy and colleagues agreed that the therapeutic relationship activated the patient's expectations about attachment relationships. However, he did not believe that the benefits to the patient of therapy stemmed from an increase in attachment security.

Instead, Fonagy praised Patricia Crittenden for the proposal that, when the attachment system is activated by perceived external or internal sources of threat, both children and adults would be predisposed to fall back on automatic processing.⁸⁴ Drawing ideas from Linda Mayes, Fonagy and Bateman redescribed this process as a tendency among patients with BPD to fall into an implicit, non-mentalizing form of processing when the attachment system was activated.⁸⁵ As the attachment bond between therapist and client intensifies, the quality of mentalizing of patients, especially those with BPD, was anticipated to drop off. The therapeutic task, then, was to help patients retain their capacity to mentalize even in the context of attachment relationships. It was anticipated that successes in the mentalization of attachment relationships, because of their foundational position for psychological development, would be especially likely to be extrapolated to other areas of life:

⁸⁴ Fonagy, P. (1998). 'Moments of Change in Psychoanalytic Theory: Discussion of a New Theory of Psychic Change'. *Infant Mental Health Journal*, 19(3): 346–353: 'At the heart of this new theory is the notion of implicit or procedural memory borrowed from cognitive science (Schacter, 1992). Bob Clyman (1991) should be credited with bringing this idea to the attention of psychoanalysts, and Crittenden (1990) with integrating the idea with attachment theory' (p. 348).

⁸⁵ Fonagy, P. and Bateman, A. W. (2006). 'Mechanisms of Change in Mentalization-Based Treatment of BPD'. *Journal of Clinical Psychology*, 62(4): 411–430. See also Mayes, L. C. (2006). 'Arousal Regulation, Emotional Flexibility, Medial Amygdala Function, and the Impact of Early Experience: Comments on the Paper of Lewis et al'. *Annals of the New York Academy of Sciences*, 1094: 178–192.

For reasons that may have evolutionary significance, the intense activation of the networks underpinning attachment feelings and experiences also appears to inhibit the intensity of cognitive and emotional scrutiny over mental contents. This state of affairs creates a unique opportunity for the psychotherapist. By balancing the activation of attachment against the presentation of negative mental contents, they are able to present new stimuli (mental contents) to the patient without evoking mental resistance against the incorporation of new ways of experiencing the world into existing cognitive–emotional schemata.⁸⁶

Fonagy and colleagues proposed that a therapeutic focus on mentalization would be of benefit to patients with most mental health problems. Putting matters stridently, Fonagy and Target stated that, ‘in our view, change can happen *solely* through the revival of reflective function.’⁸⁷ However, they argued that a focus on mentalization in the context of attachment relationships would especially benefit patients with BPD, because they regarded this condition as primarily reflecting deficits in mentalizing in the context of attachment relationships. If difficulties with mentalizing could be drawn as a spectrum, BPD was conceptualized as the extreme end of this spectrum. However, at times, they also described the pathway to other conditions—such as dissociative disorders—in exactly the same terms as the pathway to BPD, in terms of a turning away from conceiving of the mind of the caregiver resulting in mentalization deficits.⁸⁸ By the early 2000s, such diffuse appeal to the same explanatory mechanism and psychological constructs was starting to raise serious questions about the exact meaning of mentalization and about the specificity of the developmental model underpinning it. However, this model was anchored in place, at least for a time, by the manner in which Fonagy and colleagues made use of the concept of disorganized attachment.

Disorganization and mentalization

The concept of disorganized attachment was introduced by Mary Main and Judith Solomon, and technically refers to the display of conflicted, confused, or apprehensive behaviour by young children in the Strange Situation procedure. In writing with the Steeles in the early 1990s, reporting empirical findings from studies using the Strange Situation, this was the meaning of the term in Fonagy’s work.⁸⁹ Fonagy and Target were enthusiastic, stating that

⁸⁶ Fonagy, P. and Adshead, G. (2012). ‘How Mentalisation Changes the Mind’. *Advances in Psychiatric Treatment*, 18(5): 353–362, p. 359.

⁸⁷ Fonagy, P. and Target, M. (2005). ‘Some Reflections on the Therapeutic Action of Psychoanalytic Therapy’, in J. Auerbach, K. Levy, and C. E. Shaffer (eds), *Relatedness, Self-definition and Mental Representation: Essays in Honor of Sidney J. Blatt*, New York: Taylor & Francis, pp. 191–212, p. 200, italics added.

⁸⁸ E.g. Fonagy, P. ([1997] 2002). ‘Multiple Voices versus Meta-Cognition: An Attachment Theory Perspective’, in V. Sinason (ed.), *Attachment, Trauma and Multiplicity: Working with Dissociative Identity Disorder*, London: Brunner-Routledge, pp. 71–85. The in-but-mostly-out status of dissociation within Fonagy’s explanatory model would later be a point of acknowledged self-criticism e.g. Ensink, K., Bégin, M., Normandin, N., Godbout, N., and Fonagy, P. (2017). ‘Mentalization and Dissociation in the Context of Trauma: Implications for Child Psychopathology’. *Journal of Trauma & Dissociation*, 18(1): 11–30: ‘The finding that dissociation is central in processes associated with externalizing and sexualizing behaviors, partly by itself and partly with RE, nuanced our predictions based on the model proposed by Bateman and Fonagy (2008) in which mentalizing is considered to be the key mental process that has a social regulation role ... dissociation was not considered in Bateman and Fonagy’s conceptual model’ (p. 24).

⁸⁹ However, already in these writings, non-apprehensive conflict was downplayed in favour of an emphasis on fear and confusion. See e.g. Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). ‘The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment’. *Infant Mental Health Journal*, 12(3): 201–218: ‘A small group of infants show a fourth pattern of response: one of confusion and disorganization (D)’ (p. 206).

‘the most promising area of attachment research is undoubtedly the study of disorganized/disoriented attachment behaviour’.⁹⁰ However, in the context of this enthusiasm, the term ‘disorganized attachment’ took on a further life through the 1990s and 2000 in the writings of Fonagy and his immediate collaborators. This usage was cut free from the actual observable characteristics of disorganized attachment. The most important factor was that Main, Solomon, and colleagues had used the term ‘disorganized’ to refer to their observations, but had intended a technical meaning that they did too little to clarify for the reader: observations of conflict, confusion, and apprehension, from which can be inferred conflict of expectations about the availability of the caregiver as a safe haven. Seeing the confusion their early writings have caused for subsequent researchers, Main, Solomon, and colleagues have accepted a portion of the responsibility for subsequent hazy or spurious uses of the category.⁹¹

Fonagy’s group, however, were especially vulnerable to this wider problem. After Howard and Miriam Steele left for New York, none of Fonagy’s circle of collaborators were reliable coders of disorganized attachment, allowing their theorizing to drift free from the operationalization of the construct.⁹² Use of the concept of disorganized attachment seems to have been implicitly shaped by the narrative need for an antagonist within stories about the development of mentalizing.⁹³ This was likely unintentional. However, unmoored to the construct’s operationalization, the result has been an account of disorganization that slips into metaphor, or—since the metaphorical status is not signalled—into ‘pretend mode’ (see Chapter 5).

To take one example: in a chapter from 2010, Fonagy, Luyten, Bateman, Gergely, Strathearn, Target, and Allison depict disorganized attachment as ‘undirected/bizarre behaviour’ and ‘trying to escape the situation even in the caregiver’s presence’.⁹⁴ However

⁹⁰ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 244.

⁹¹ Duschinsky, R. and Solomon, J. (2017). ‘Infant Disorganized Attachment: Clarifying Levels of Analysis’. *Clinical Child Psychology and Psychiatry*, 22(4): 524–538; Reijman, S., Foster, S., and Duschinsky, R. (2018). ‘The Infant Disorganised Attachment Classification: “Patterning within the Disturbance of Coherence”’. *Social Science & Medicine*, 200: 52–58. On the ‘concept creep’ of ideas like disorganized attachment that signify psychological harm, see also Haslam, N. (2016). ‘Concept Creep: Psychology’s Expanding Concepts of Harm and Pathology’. *Psychological Inquiry*, 27(1): 1–17.

⁹² Target, M., Mayes, L., and Bach, S. (2000). ‘Panel 4: The Pathology of the Self: The Fragmented Self, Disorders of the Self, and the Dissolution of the Self’. *Journal of Infant, Child and Adolescent Psychotherapy*, 1(3): 63–72. Hepworth remarked that ‘it is interesting to note Breuer and Freud’s (1895) comment in *Studies on Hysteria*: “It is easy to fall into a habit of thought which assumes that every substantive has substance behind it. We find as time goes on, that we have actually formed an idea which has lost its metaphorical nature, and which we can manipulate easily, as though it were real” (pp. 227–228) (p. 63). See also Reijman, S., Foster, S., and Duschinsky, R. (2018). ‘The Infant Disorganised Attachment Classification: “Patterning within the Disturbance of Coherence”’. *Social Science & Medicine*, 200: 52–58.

⁹³ E.g. Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). *Affect Regulation, Mentalization, and the Development of the Self*, London: Karnac Books: ‘At the extreme end of the safety-to-fearfulness dimension, there can be no strategy because the attachment system is not there to sustain a consistent set of defences. In these cases the interpretive mechanism that sustains social relations functions so poorly that the capacity to arrive at representations of the motivational or epistemic mind states of the other, independent of those of the self, is profoundly compromised. This is attachment disorganisation or, rather, the absence of the mental function that sustains attachment. Thus we conceive of attachment disorganisation as lying at the opposite end of the scale to attachment security and as an indicator of the regular failure of the interpersonal interpretive mechanism’ (pp. 135–136). Fonagy has subsequently acknowledged that the 2002 book was ‘patchy’. One of the most powerful contributing problems is the multiple and conflicting ways that the different chapters use the concepts of disorganization and mentalizing. BBC Radio 4 (2020). ‘Peter Fonagy on a Revolution in Mental Health Care’. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>.

⁹⁴ Fonagy, P., Luyten, P., Bateman, A. W., Gergely, G., Strathearn, L., Target, M., and Allison, E. (2010). ‘Attachment and Personality Pathology’, in J. Clarkin, P. Fonagy, and G. Gabbard (eds), *Psychodynamic Psychotherapy for Personality Disorders: A Clinical Handbook*, Arlington, VA: American Psychiatric Publishing, Inc., pp. 37–88, 40–41.

'undirected/bizarre behaviour' is a partial description of one of the seven indices of disorganized attachment ('undirected/misdirected behaviour'). Trying to escape the situation is actually coded as avoidant attachment ('5b' in the avoidance scale protocols: baby 'tries to go past her out the door'), and is not necessarily treated by coders as disorganization.⁹⁵ Fonagy and colleagues then say, 'in these individuals, the attachment system may be quite readily triggered, and they may appear to be constantly preoccupied with attachment relationships'.⁹⁶ This is ambivalent/resistant, not disorganized attachment. Some infants in dyads classified as disorganized show preoccupation with attachment relationships. Others combine conflicted, confused, or apprehensive behaviours with avoidance of the caregiver. Others combine these behaviours with otherwise secure use of the caregiver as a secure base and safe haven.

Over time, Fonagy has come to increasingly acknowledge that the relationship between disorganized attachment and mentalizing is weak and probably indirect (see Chapter 7).⁹⁷ However, through the 1990s, 2000s, and early 2010s, reference to disorganized attachment in causal accounts of disturbed development by Fonagy and colleagues was pervasive,⁹⁸ and one of the least meticulous aspects of their developmental model. Rather than the observable behaviour, or known correlates, the characterization of disorganized attachment seemed shaped more by connotations of the word 'disorganized' itself, which in ordinary language means to be thrown into chaos. This was facilitated, not just in the work of Fonagy and colleagues but also in much other speculation about disorganized attachment, by Mary Main's introduction of the term with a technical meaning, but without adequate definition or explanation.⁹⁹ In a chapter from 2007, Fonagy and Higgitt acknowledged this problem, observing that 'disorganised attachment is currently poorly described in the literature,' implying that this has hampered their own use of the concept.¹⁰⁰

Examination of appeals to disorganized attachment in their writings reveals that Fonagy and colleagues have offered diverse non-overlapping descriptions of the nature of

⁹⁵ Ainsworth, M., Blehar, M., Waters, E., and Wall, S. ([1978] 2015). *Patterns of Attachment: A Psychological Study of the Strange Situation*, Bristol: Psychology Press, p. 347.

⁹⁶ Fonagy, P., Luyten, P., Bateman, A. W., Gergely, G., Strathearn, L., Target, M., and Allison, E. (2010). 'Attachment and Personality Pathology', in J. Clarkin, P. Fonagy, and G. Gabbard (eds), *Psychodynamic Psychotherapy for Personality Disorders: A Clinical Handbook*, Arlington, VA: American Psychiatric Publishing, Inc., pp. 37–88, 40–41.

⁹⁷ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds about: Part 1. Borderline Personality Disorder as a Limitation of Resilience'. *Borderline Personality Disorder and Emotion Dysregulation*, 4: 11: 'We have previously placed considerable weight on the nature of attachment disorganization in our accounts of BPD based on the mentalizing model. We maintain that the role of attachment is highly significant in the developmental origins of PD. However, we argue that its role might perhaps be best understood as only one (albeit very important) form of content learned from the social environment. This is congruent with recent work suggesting that the relationship between infant attachment status and later outcomes is more complicated than that suggested by early attachment studies. Other findings have suggested limited evidence for linking childrearing environments to later outcomes.'

⁹⁸ Fonagy was also a strong advocate for the category to professional groups, above all clinicians and teachers: e.g. speaking to the House of Commons Education Committee (2016) 'Mental Health and Well-Being of Looked After Children', HC 481, Wednesday, 13 January. Accessed at: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/education-committee/mental-health-and-wellbeing-of-looked-after-children/oral/26927.pdf>: 'Here is agreed risk and also an opportunity for teachers to benefit from understanding problems that arise in relation to disrupted attachment and the disorganisation of attachment. In the NICE [National Institute for Health and Care Excellence] guidance development group that I chaired, one of the recommendations that we identified was in relation to training teachers to be more aware of the attachment issues.'

⁹⁹ Duschinsky, R. (2015). 'The Emergence of the Disorganised/Disoriented (D) Attachment Classification, 1979–1982'. *History of Psychology*, 18(1): 32–46.

¹⁰⁰ Fonagy, P. and Higgitt, A. (2007). 'The Early Social and Emotional Determinants of Inequalities in Health', in G. Baruch, P. Fonagy, and D. Robins (eds), *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, Chichester: John Wiley & Sons, pp. 3–34, p. 13.

Table 3.1 Descriptions of disorganized attachment in Fonagy and colleagues

	Characterization	Example
1	A combination of avoidant and resistant attachment behaviour	Luyten, P., van Houdenhove, B., Lemma, A., Target, M., and Fonagy, P. (2012). 'A Mentalization-Based Approach to the Understanding and Treatment of Functional Somatic Disorders.' <i>Psychoanalytic Psychotherapy</i> , 26(2): 121–140: 'When attachment figures have been insufficiently available, abusing and/or nonresponsive, attachment deactivating or hyperactivating strategies, or a combination of both, as is typically observed in disorganized attachment, become a habitual response to stress' (p. 126).
2	The breakdown of combinations of avoidant and resistant attachment behaviour into a diffuse state	Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression,' in A. W. Bateman and P. Fonagy (eds.), <i>Handbook of Mentalizing in Mental Health Practice</i> (pp. 385–418). Washington, DC: American Psychiatric Association: 'Patients with dysfunctions of the attachment system may show a sequence, moving from deactivating strategies to hyperactivating strategies if the former fail, or may show marked oscillations between the two strategies ... depressed individuals with these features often show many borderline features ... At higher levels of personality functioning, these patients may show a pattern of fearful-avoidant attachment, and thus a typical approach-avoidance conflict with regard to attachment relationships. This conflict is typically expressed either in very brief and barren attachment relationships, as in patients with schizoid and schizotypal features, or in hypermentalising accounts revealing both their desire for and their fear of relationships, as in patients with hysterical features. It is our impression that in individuals at lower levels of personality functioning, disorganised attachment is related to identity diffusion, while individuals at higher levels of personality functioning seem to be characterised by marked rigidity. Hence, whereas the former group may be particularly responsive to mentalisation interventions, the latter may show high levels of pseudomentalising, which is not only difficult to distinguish from genuine mentalisation, but also perhaps more difficult to treat' (pp. 413–414).
3	A hyperactivating strategy similar to anxious attachment, though often with 'frantic attempts to downregulate'	Fonagy, P. (1998). 'Prevention, the Appropriate Target of Infant Psychotherapy.' <i>Infant Mental Health Journal</i> , 19(2): 124–150: 'Anxious attachment or disorganization of attachment develops (Erickson, Sroufe, and Egeland, 1985; Greenberg et al., 1993) to the mother as part of a strategy to avoid being blocked from access to her' (p. 135). Bateman, A. W. and Fonagy, P. (2016). <i>Mentalization-Based Treatment for Personality Disorders: A Practical Guide</i> (2nd edn), Oxford: Oxford University Press: 'Disorganised—Incoherent: hyperresponsive, but often frantic attempts to downregulate' (p. 123).
4	The absence of coherent affect regulation	Fonagy, P. and Higgitt, A. (2000). 'Early Influences on Development and Social Inequalities: An Attachment Theory Perspective,' in A. R. Tarlov and R. F. S. Peter (eds), <i>The Society and Population Health Reader. Vol. II: A State and Community Perspective</i> (pp. 104–130), New York: New Press: 'Disorganised or unresolved patterns are held to be indicative of an absence of coherent affect-regulation strategies' (p. 109).

Table 3.1 Continued

	Characterization	Example
5	The absence of any strategy to interpersonal relationships—for instance, the absence of help-seeking in the context of need	Stein, H., Koontz, A. D., Fonagy, P., Allen, J. G., Fultz, J., Brethour, J. R., ... and Evans, R. B. (2002). 'Adult attachment: What are the underlying dimensions?'. <i>Psychology and Psychotherapy: Theory, Research and Practice</i> , 75(1): 77–91: 'With more insecurity comes an increased need for a strategy to manage attachment relationships. At the extreme insecure end of the dimension, adaptive strategies again diminish, from dismissing or preoccupied strategies to no coherent strategy. Thus, there are probably three important positions, conflated until now: low insecurity with no need for special strategy; moderate to high insecurity with specific strategies for sustaining relationships; and extreme insecurity with a disorganized attachment system that precludes strategy' (p. 88). Bevington, D., Fuggle, P., and Fonagy, P. (2015). 'Applying Attachment Theory to Effective Practice with Hard-To-Reach Youth: The AMBIT Approach', <i>Attachment & Human Development</i> , 17(2): 157–174: 'The core assumption of AMBIT is that those considered "hard to reach" are hard to reach for reasons: their avoidance of help is frequently active and intentional, rooted as it may be in profound disorganizations within their attachment systems' (p. 160).

disorganized attachment (see Tables 3.1 and 3.2). Each of these five descriptions is an extrapolation from *some* among the behaviours used to code disorganized attachment. The most common description of disorganized attachment as a whole, especially in writings with Luyten, is of a combination of avoidant and resistant attachment.¹⁰¹ In fact, only a small fraction of child–caregiver dyads who receive a disorganized attachment classification show both avoidance and resistance.¹⁰² Another common description is of disorganized

¹⁰¹ This claim seems to be based on a conflation of disorganized attachment as an assessment for the Strange Situation with the co-occurrence of anxiety and avoidance in self-report measures of attachment. In the early 1990s, Shaver and colleagues claimed that the co-occurrence of attachment anxiety and avoidance is analogous to Main and Solomon's disorganized attachment classification. This was based on a proposal by Bartholomew, who called this co-occurrence 'fearful attachment'. However, examination of Bartholomew's 1989 unpublished doctoral thesis indicates an important qualification: Bartholomew regarded fearful attachment not as a general analogue for the disorganized attachment classification, but as an analogue for approach–avoidance conflict specifically. Jacobvitz and colleagues criticized the assumption by Shaver and others that disorganized attachment is the same as co-occurrence of attachment anxiety and avoidance. Brennan, K. A., Shaver, P. R., and Tobey, A. E. (1991). 'Attachment Styles, Gender, and Parental Problem Drinking'. *Journal of Social and Personal Relationships*, 8: 451–466; Bartholomew, K. (1989). *Attachment Styles in Young Adults: Implications for Self-Concept and Interpersonal Functioning*, unpublished doctoral dissertation, Stanford University; Jacobvitz, D., Curran, M., and Moller, N. (2002). 'Measurement of Adult Attachment: The Place of Self-Report and Interview Methodologies'. *Attachment & Human Development*, 4(2): 207–215, p. 209. There is also no prospective or concurrent association between measures of disorganized attachment and co-occurrence of anxiety and avoidance in self-report measures of attachment. Roisman, G. I., Holland, A., Fortuna, K., Fraley, R. C., Clausell, E., and Clarke, A. (2007). 'The Adult Attachment Interview and Self-Reports of Attachment Style: An Empirical Rapprochement'. *Journal of Personality and Social Psychology*, 92(4): 678–697; Fraley, R. C., Roisman, G. I., Booth-LaForce, C., Owen, M. T., and Holland, A. S. (2013). 'Interpersonal and Genetic Origins of Adult Attachment Styles: A Longitudinal Study from Infancy to Early Adulthood'. *Journal of Personality and Social Psychology*, 104(5): 817–838, web-based supplement C.

¹⁰² Some infants show co-occurrence of avoidance and resistant attachment, and on this basis receive a disorganized classification. But this is hardly the predominant phenotype. A rare form of behaviour under Index I of the Main and Solomon coding protocols is hyperarousal followed by relatively abrupt downregulation. However, this is by no means a behaviour representative of the category as a whole. Most child–caregiver dyads who receive a disorganized classification do so on the basis of other forms of behaviour in the Main and Solomon indices, such as other kinds of approach/avoidance conflict, stereotypes or confusion, or apprehension without any apparent conflict. The predominance of different forms is anticipated to depend somewhat on the kinds of adversities faced by the sample. Solomon, J., Duschinsky, R., Bakkum, L., and Schuengel, C. (2017). 'Toward an Architecture of Attachment Disorganization: John Bowlby's Published and Unpublished Reflections'. *Clinical Child Psychology and Psychiatry*, 22(4): 539–560.

Table 3.2 Descriptions of the proximal mechanism of disorganized attachment in Fonagy and colleagues

	Characterization	Example
1	Principally driven by the child's constitutional characteristics	Fonagy, P., Bateman, A. W., and Luyten, P. (2012). 'Introduction and Overview', in <i>Handbook of Mentalizing in Mental Health Practice</i> (pp. 3–42). Washington, DC: American Psychiatric Publishing: 'Abnormal hippocampal and hyperactive amygdala functioning in children who are at risk for later developing BPD may cause highly anxious and emotionally labile infant behaviour and thus an infant who is not able to benefit from the regulating qualities of the attachment relationship. This constellation might well create disorganised attachment relationships, principally driven by the child's constitutional characteristics' (p. 8).
2	Dissociation	Fonagy, P. (1999). 'The Transgenerational Transmission of Holocaust Trauma', <i>Attachment & Human Development</i> , 1(1): 92–114: 'The behaviour of the infant in this situation gives clear indication of multiple, incoherently integrated structures, highly reminiscent of dissociative adult patients' (p. 97).
3	Fear of the caregiver	Fonagy, P. (2003). 'The Violence in our Schools: What can A Psychoanalytically Informed Approach Contribute?' <i>Journal of Applied Psychoanalytic Studies</i> , 5(2): 223–238: 'Neither disorganization of the attachment system nor violence are static. The attachment system becomes disorganized in the face of external and internal demands to deal with safety issues. It is the persistence of fear alongside the activation of the attachment system that accounts for the picture of disorganization. This means that assessment must always be focused on the interplay of the situation and the person. It is neither the situation nor the person that represents the risk, but rather the interaction of the two' (p. 227).
4	Disruption of the 'self'	Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). <i>Affect Regulation, Mentalization, and the Development of the Self</i> . London: Karnac Books: 'The disorganisation of self disorganises attachment relationships by creating a constant need for this projective identification' (p. 12).
5	A preoccupation with ensuring perfect contingency from caregivers	Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). <i>Affect Regulation, Mentalization, and the Development of the Self</i> . London: Karnac Books: 'An early association between disorganised infant attachment and an abnormal preference for perfect contingencies' (p. 250).

attachment as the absence of any logic to interpersonal interactions in the context of emotional dysregulation. Again, there are some infants in the disorganized category who show pervasive disturbances of affective control. However, these are a vast minority.¹⁰³ Children classified as disorganized under the Main and Solomon protocols do not show a coherent avoidant or resistant strategy; but this does not mean that their behaviour lacks strategy or

¹⁰³ Lyons-Ruth and colleagues have criticized the characterization of disorganized attachment as the absence of any strategy to interpersonal relationships, observing that a majority of infants who receive the classification nonetheless show proximity-seeking and contact maintenance with their caregiver when distressed, in the manner of the secure category. Lyons-Ruth, K., Bureau, J. F., Easterbrooks, M. A., Obsuth, I., Hennighausen, K., and Vulliez-Coady, L. (2013). 'Parsing the Construct of Maternal Insensitivity: Distinct Longitudinal Pathways Associated with Early Maternal Withdrawal'. *Attachment & Human Development*, 15(5–6): 562–582.

any form of regulation.¹⁰⁴ Part of the problem is that it is not entirely clear what is meant by 'absence of coherent affect regulation' or 'absence of any strategy'. But, in general terms, they are at best very partial characterizations of disorganized attachment, and rather misleading. It would appear that Fonagy and colleagues are not primarily talking about the same thing that developmental attachment researchers have operationalized under the rubric of 'disorganized attachment'. This is not to minimize the importance of the process they are discussing. However, it remains insufficiently clear what this process is, and use of the term 'disorganized attachment' is misleading.

As well as offering descriptions of disorganized attachment at a behavioural level, Fonagy and colleagues variously offer five different characterizations of the proximal mechanism of disorganized attachment: constitutional factors; dissociation; fear of the caregiver; disruption of the 'self'; and a preoccupation with perfect contingencies in attachment relationships. It is absolutely possible that the forms of behaviour observed by Main and Solomon in the Strange Situation are prompted by these diverse mechanisms. It may be that all these pathways lead to a single phenomenon of 'disorganized attachment'; or they may be different pathways to different forms of confusion, conflict, or apprehension in attachment relationships. The greater problems, however, are that Fonagy and colleagues do not reconcile these different accounts theoretically, and some are lacking much in the way of evidence.¹⁰⁵ For instance, in 2002, Fonagy and colleagues write that 'The disorganization of self disorganizes attachment relationships'; in 2004, they write 'disorganized attachment is rooted in a disorganized self'.¹⁰⁶ Of course, it is quite possible for there to be a bidirectional relationship. However, the two claims are not reconciled. A hazy notion of disorganized attachment, combined with a hazy concept of 'self' (see Chapter 6), appears to have hindered the precision of causal claims, and the coherence of the theory built from them. In their writings, Fonagy and colleagues also offer various characterizations of what disorganized attachment entails at a cognitive level. These include the 'absence of stable representation of interactions and therefore predictability', internalization of 'confusing and hostile parenting',¹⁰⁷ and failure to develop emotional awareness.¹⁰⁸ Again, however, these accounts are not reconciled.

¹⁰⁴ For instance, a toddler who is solicitous and caring towards their parent when distressed is not showing an avoidant or resistant strategy, and would likely be placed in the disorganized classification, despite the fact that their behaviour is strategic and their affects may be coherent in their expression (Chapter 7). Mary Main and Erik Hesse, personal communication, August 2019: 'Development allows humans to override a behavioural system in other ways than the two conditional strategies, producing a wider variety of potential strategies than those available to infants. These might well not be conditional strategies in the technical sense of being a behavioural repertoire made available by human evolutionary history for solving problems of survival and reproduction. They could be described as "strategic" in the non-technical sense—but it depends on how the word is being used.'

¹⁰⁵ For instance: the claim that all disorganized attachment represents a preoccupation with perfect contingencies in attachment relationships, emphasized especially by Fonagy and colleagues in the early 2000s, seems to have been downplayed over the past decade. One reason is that the evidence base for the claim has not expanded beyond Gergely's original studies. The primary subsequent source of supportive evidence comes from the observation of poor caregiver–infant contingency among infants who would later receive a disorganized classification in the Strange Situation by Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., . . . and Feldstein, S. (2010). 'The Origins of 12-Month Attachment: A Microanalysis of 4-Month Mother–Infant Interaction.' *Attachment & Human Development*, 12(1–2): 3–141. However, there would be a variety of other ways of interpreting this finding, besides the idea that all disorganization reflects preoccupation with perfect contingencies. See e.g. Bernier, A. and Meins, E. (2008). 'A Threshold Approach to Understanding the Origins of Attachment Disorganization.' *Developmental Psychology*, 44(4): 969–982; van IJzendoorn, M. H. and Bakermans-Kranenburg, M. J. (2019). 'Bridges across the Intergenerational Transmission of Attachment Gap.' *Current Opinion in Psychology*, 25: 31–36.

¹⁰⁶ Bateman, A. W. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorders: Mentalization Based Treatment*, Oxford: Oxford University Press, p. 90.

¹⁰⁷ Fonagy, P. (1999). 'Points of Contact and Divergence between Psychoanalytic and Attachment Theories: Is Psychoanalytic Theory Truly Different?' *Psychoanalytic Inquiry*, 19(4): 448–480, p. 467.

¹⁰⁸ Fonagy, P. and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology.' *Journal of Personality Disorders*, 20: 544–576: 'Disorganised infants, even if interpersonally perceptive, fail to integrate this emotional awareness with their self-organisation' (p. 565).

Furthermore, this picture seems to have been based in part on the connotations of the word ‘disorganized’. The primary trainers in coding disorganized attachment, Alan Sroufe and Elizabeth Carlson, have criticized the depiction of disorganized attachment as lack of predictability, which they regard as an erroneous extrapolation from the fact that some, not all, children in this category show confused behaviours.^{109,110} The idea that disorganized attachment entails confusion of representations about attachment figures may also have partly have stemmed from Main’s original Berkeley study, where some children showing disorganized attachment in infancy showed themes of conflict, confusion, and/or fear in story-stem narratives about attachment relationships at age 6.¹¹¹ However, this is not necessarily the same as the absence of stable representations about interactions. Nor does it imply that the children lack predictability in their behaviour. In fact, in their classic study, Main and Cassidy found that infants from dyads classified as disorganized showed highly concerted, goal-directed behaviour on reunion at age 6 with their caregivers, either controlling and solicitous or controlling and punitive.¹¹²

As we will see in the subsequent chapters, many of the themes gestured to by disorganized attachment—fear, constitutional factors, dissociation, disruption of the ‘self’—have played a significant role in the developing conceptual system of Fonagy and colleagues. Yet, recognition of the distinct role of these factors has been hindered by the absorptive construct of disorganized attachment. The concept of disorganized attachment, as a container, has tended to devour the elements that it ostensibly contains. This problem has been supported by the lack of clarity in attachment research in general, and in the work of Fonagy and colleagues in particular, regarding what exactly constitutes an attachment relationship as opposed to other kinds of relationship. As articulation of role of fear, constitutional factors, dissociation, disruptions to the ‘self’, and other more specific factors has improved in the work of Fonagy and colleagues over the decades, appeal to disorganized attachment by Fonagy and colleagues has been dropping away in recent years. These conceptual improvements stem especially from self-criticism and important theoretical innovations in the 2000s, prompting revisions in the model used by Fonagy and colleagues in conceptualizing the development of mentalizing and non-mentalizing. These revisions will be the focus of the next chapter.

¹⁰⁹ Padrón, E., Carlson, E., and Sroufe, A. (2014). ‘Frightened Versus Not Frightened Disorganized Infant Attachment: Newborn Characteristics and Maternal Caregiving’. *American Journal of Orthopsychiatry*, 84(2): 201–208.

¹¹⁰ See also Duschinsky, R. (2018). ‘Disorganization, fear and attachment: Working towards clarification’. *Infant Mental Health Journal*, 39(1): 17–29.

¹¹¹ Kaplan, N. (1987). *Individual Differences in 6-Years Olds’ Thoughts about Separation: Predicted from Attachment to Mother at Age 1*. Unpublished doctoral dissertation, Berkeley, CA: Department of Psychology, University of California.

¹¹² Main, M. and Cassidy, J. (1988). ‘Categories of Response to Reunion with the Parent at Age 6’. *Developmental Psychology*, 24(3): 415–426. It should also be noted that subsequent findings since Main have found much less stability than the original study between disorganized attachment in the Strange Situation and later representational measures ($r = .26$). Pinquart, M., Feußner, C., and Ahnert, L. (2013). ‘Meta-Analytic Evidence for Stability in Attachments from Infancy to Early Adulthood’. *Attachment & Human Development*, 15(2): 189–218. An important reason for this difference was that Main and colleagues developed their story-stem coding system quasi-inductively on the sample, and then examined its retrodiction of the Strange Situation. Associations between the Strange Situation and the story-stem narratives at age 6 was therefore not an independent finding but a methodological principle for the elaboration of the 6-year coding system. It should not be assumed that disorganized attachment can be treated as meaning the absence of stable representation of interactions and therefore predictability.

4

Mentalization in transition

Introduction

By the end of the 1990s, Fonagy and colleagues had worked for 10 years in developing the concept of mentalization, alongside a developmental model of its emergence from early child-caregiver relationships. They had highlighted the importance of mentalization for psychotherapy, and especially for work with patients with borderline personality disorder (BPD). Furthermore, the reflective functioning scale for the Adult Attachment Interview had been validated as a measure of mentalization in the context of attachment relationships and demonstrated impressive prediction to infant-caregiver relationships as assessed in the Strange Situation procedure. Yet, over the 2000s, Fonagy and colleagues became more and more dissatisfied with their initial model, identifying that the concept of 'mentalization' was too absorptive and that they had placed too much weight on early childhood experiences and the role of disorganized attachment. Attempts to correct these problems led to a revised and more mature account of forms of mentalizing.¹ In this chapter, we will trace these developments of the 2000s, paying particular attention to Fonagy and Luyten's 2009 account of four dimensions of mentalizing. The chapter will close by attempting to discern the underpinning logic of the concept of mentalization across its diverse uses and definitions. Synthesising the various uses of the term, mentalization will be ultimately defined as a capacity to:

1. conceive of
2. and make available for reconsideration
3. the thoughts
4. and feelings
5. implicated in motivations and intentions
6. in order to account for and explain
7. the observable social behaviour
8. and present and past perceptual experience
9. of oneself
10. and others.

Self-criticism

As we saw in the previous chapter, in the 1990s, Fonagy and colleagues had tended to write about reflective functioning and mentalization as if they were relatively basic and discrete.

¹ The characterization of the 2009 paper as the commencement of a 'mature model of mentalizing' is supported by Fonagy's description of the work in the UCL REF 2014 impact case study as the place where 'the model is fully developed': University College London (2014). REF 2014: Psychology, Psychiatry and Neuroscience impact case study. Accessed at: [https://results.ref.ac.uk/\(S\(jj2mvvb3fbee3zpqblartx2d\)\)/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202](https://results.ref.ac.uk/(S(jj2mvvb3fbee3zpqblartx2d))/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202).

However, in the late 1990s and especially from the early 2000s, the terms ‘reflective functioning’ and ‘mentalizing’ were increasingly recognized as descriptive accounts of composite psychological processes. Particular attention was given to effortful attentional control and affect regulation as developmental requisites² and component elements of mentalization, as well as—additionally—effects of mentalization. For instance, attentional control would help an individual retain access to inferential information about the other’s mind even in the context of distractions or stress. In turn, awareness of the other’s mind could help inform attentional processes, helping to organize and scaffold them.³

Fonagy characterized attentional control, affect regulation, and mentalizing as interacting elements of a meta-process he termed the ‘interpersonal interpretive function.’⁴ This was the beginning of the network approach to mental health that would be developed in the work of Fonagy and colleagues in later years (see Chapter 7). In the late 1990s and early 2000s, however, Fonagy and colleagues initially adopted a more linear model, grounded in attachment theory. They proposed the importance of attachment patterns for attentional control and affect regulation, which in turn serve as the scaffolding for mentalization in general.⁵

Despite this importance given to attachment, a transition in Fonagy and colleagues’ perspective was growing acknowledgement that adverse early caregiving experiences was not the only factor in play. Already in 1997, Fonagy and Target acknowledged that ‘the child’s biological vulnerabilities such as hyperactivity, attention problems, low impulse control, are all likely to obstruct the opportunity the child has for evolving a mentalized reflective model.’⁶ One early project addressing potential biological vulnerabilities was Pasco Fearon’s 1998 doctoral study under Fonagy’s supervision. Sixty-two same-sex twins were seen in the Strange Situation procedure with a primary caregiver. Fearon found that there was no more concordance in the twins’ patterns of attachment to a primary caregiver than would be expected by chance.⁷ These findings were then replicated in collaboration with van IJzendoorn

² Fonagy, P. and Target, M. (2002). ‘Early Intervention and the Development of Self-Regulation.’ *Psychoanalytic Inquiry*, 22(3): 307–335: ‘There is good reason to believe that the capacity for effortful attentional control may be necessary for the development of the mentalizing function. First, there is a chronological overlap between these developments and they share a common developmental timetable. Second, brain imaging studies suggest that the structures for mentalizing and those for effortful control are at least close if not overlapping in location. Third, individuals with Theory of Mind (ToM) deficits (e.g. individuals with autism) also have difficulty with executive function tasks that require inhibitory control. Finally, success at the mentalizing tasks requires a certain capacity for attentional control’ (p. 322).

³ See also Jones, B. and Allison, E. (2010). ‘An Integrated Theory for Attention-Deficit Hyperactivity Disorder [ADHD].’ *Psychoanalytic Psychotherapy*, 24(3): 279–295.

⁴ Fonagy, P. (2002). ‘The Internal Working Model or the Interpersonal Interpretive Function.’ *Journal of Infant, Child, and Adolescent Psychotherapy*, 2(4): 27–38.

⁵ Fonagy, P. and Target, M. (2006). ‘The Mentalization-Focused Approach To Self Pathology.’ *Journal of Personality Disorders*, 20: 544–576: ‘An enfeebled attentional control system is a likely consequence of attachment disorganisation, perhaps linked with enfeebled affect representation, and serves to undermine the development of mentalisation as well as its appropriate functioning in later development. The prepotent response is to attribute one’s own mental state to the other. Attentional control is essential if the child is to arrive at a differentiation of their own and others’ thoughts, feelings, beliefs and desires. The disruption of attentional control is likely to account for many instances where we encounter temporary and selective disruptions of mentalising’ (p. 556). See also Fonagy, P. and Target, M. (2002). ‘Early Intervention and the Development of Self-Regulation.’ *Psychoanalytic Inquiry*, 22(3): 307–335: ‘All the key mechanisms underpinning the enduring effects of early relationship experiences interface with individuals’ capacity to control (a) their reaction to stress, (b) their capacity to maintain focused attention, and (c) their capacity to interpret mental states in themselves and others’ (p. 307).

⁶ Fonagy, P. and Target, M. (1997). ‘Attachment and Reflective Function: Their Role in Self-Organization.’ *Development and Psychopathology*, 9(4): 679–700, p. 696.

⁷ Fearon, P. (1998). *Determinants of Mother-Infant Attachment Classification In Twins*. Unpublished doctoral dissertation, London: University College London.

and colleagues at Leiden.⁸ However, a later study by Fearon with 551 twin pairs, and using an adaptation of the Adult Attachment Interview for adolescents, found very substantial associations between monozygotic twins' scores for coherence and their overall security of attachment ($r = .42$), but substantially lower associations for dizygotic twins ($r = .20$).⁹ Fearon and colleagues concluded that genetic factors may have little influence on infant attachment, but that genetic expression may have greater influence by adolescence.

Fonagy and colleagues were also interested in the potential for heritable contributions to other aspects of psychological functioning relevant to mentalization, such as predisposition to mental illness. For instance, a study by Belsky, Caspi, and colleagues, in which Fonagy was involved, found that 47% of early adolescents showed clinical levels of BPD symptomatology if they had experienced maltreatment and also had a parent with mental health problem, whereas only around 7% of adolescents showed these symptoms when one or the other factor was present. Although, naturally, parental mental illness could also contribute to child caregiving and the general atmosphere of the home, the researchers interpreted these findings as indicating the potentiating role of genetic vulnerability when combined with maltreatment.¹⁰ Whereas Fonagy had earlier emphasized a solely environmental pathway to BPD, the collaboration with Belsky and Caspi indicated concern with the potential for genetic susceptibilities that could interact with environmental risk factors.

As well as growing interest in genetic factors, the 2000s also saw a strengthening emphasis in the writings of Fonagy and colleagues on the diversity of social experiences that could influence mentalizing capacities. Fonagy and Target described the importance of non-parental influence and care—for instance, the role of grandparents, older siblings, teachers, and friends: 'Children can perceive and conceive of their mental states to the extent that the behaviour of those around them has implied that they have them. This can happen through an almost unlimited set of methods.'¹¹ One important study in support of this position from the early 2000s was the trial of a school-based intervention, which will be discussed further in Chapter 9.¹² The weight given to additional discrete figures beyond the primary caregiver was in line with developments in attachment theory.¹³ Nonetheless, this innovation put additional pressure on the integrity of the conceptualization of reflective functioning and mentalizing as unitary processes.

⁸ Bokhorst, C. L., Bakermans-Kranenburg, M. J., Pasco Fearon, R. M., van IJzendoorn, M. H., Fonagy, P., and Schuengel, C. (2003). 'The Importance of Shared Environment in Mother-Infant Attachment Security: A Behavioral Genetic Study'. *Child Development*, 74(6): 1769–1782.

⁹ Fearon, P., Shmueli-Goetz, Y., Viding, E., Fonagy, P., and Plomin, R. (2014). 'Genetic and Environmental Influences on Adolescent Attachment'. *Journal of Child Psychology and Psychiatry*, 55(9): 1033–1041.

¹⁰ Belsky, D. W., Caspi, A., Arseneault, L., Bleidorn, W., Fonagy, P., Goodman, M., ... and Moffitt, T. E. (2012). 'Etiological Features of Borderline Personality Related Characteristics in a Birth Cohort of 12-Year-Old Children'. *Development and Psychopathology*, 24(1): 251–265: 'This analysis presents the number of extreme borderline group and comparison children in four groups of children defined by family history (positive/negative) and physical maltreatment (positive/negative). Children with both risk factors were 13.41 (95% CI = 8.61, 22.04) times more likely to be in the extreme borderline group than were children with neither risk factor. In contrast, children with only positive family history were only 2.53 (95% CI = 1.64, 3.92) times more likely, and children with only maltreatment were only 2.15 (95% CI = 0.69, 6.71) times more likely to be in the extreme borderline group relative to children with neither risk factor' (pp. 258–259).

¹¹ Fonagy, P., and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology'. *Journal of Personality Disorders*, 20: 544–576, p. 562.

¹² Twemlow, S. W., Fonagy, P., and Sacco, F. C. (2005). 'A Developmental Approach to Mentalizing Communities: II. The Peaceful Schools Experiment'. *Bulletin of the Menninger Clinic*, 69(4): 282–304.

¹³ van IJzendoorn, M. H. (2005). 'Attachment in Social Networks: Toward an Evolutionary Social Network Model'. *Human Development*, 48(1–2): 85–88.

In their 2003 book, *Psychoanalytic Theories*, Fonagy and Target reviewed each psychoanalytic model, and appraised its strengths and weaknesses. Admirably, their own work was not spared such appraisal. One major criticism they levelled at themselves was that ‘some people seem to be able to mentalise well but show poor functioning’. That is to say, some people’s difficulties in mentalizing primarily occur in relation to attachment relationships and not in relation to other relationships. Ultimately, they acknowledged, it was not clear that mentalization and reflective function should be regarded as ‘a single mechanism.’¹⁴ Similarly, an individual may be good at identifying mental states, such as the difference between sadness and tiredness, but not thinking about or contextualizing these states—or vice versa. The individual may be good at thinking about others, but not themselves—or vice versa. An individual may be good at sustaining mentalization but find it difficult to regain when lost—or vice versa. These are all highly clinically relevant distinctions. Over the 2000s, there was growing awareness among Fonagy and colleagues that a whole menagerie of various clinically-relevant differences could be observed within, and to an extent hidden by, the concept of mentalization if one opened up the roof and looked inside.

Impetus to address these problems came with public criticisms by Choi-Kain and Gunderson in their 2008 review of mentalization as a concept and psychological construct. Choi-Kain and Gunderson expressed concern that Fonagy and colleagues too often had veered towards treating mentalization as a single thing, which they felt was encapsulated in the way that reflective function had been operationalized:

There are limitations in interpreting the meaning of a given reflective functioning score. The capacity that is assessed by the Reflective Functioning Scale is multidimensional, with factors such as plausibility, consistency, complexity, and originality. However, the grading is done using a unidimensional score that cannot be submitted to factor analysis. In two different transcripts coded for reflective functioning with a score of 3, one transcript may reflect a consistently superficial, clichéd, and general understanding of mental states, while the other transcript reflects a highly variable capacity to understand mental states with some moments of antireflectiveness and other moments of marked reflectiveness.¹⁵

This concern was fully accepted by Target, Fonagy, and Luyten in papers over the subsequent months.¹⁶ They proposed that mentalization represented a hierarchically organized set of capacities. This meant that different components tended to occur together, giving

¹⁴ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 281. A finding that may have influenced this conclusion was that avoidant/dismissing adolescents were found to be ‘capable of mentalising in a nonpersonal context, such as the stories used to elicit mentalising, but show more difficulty in generating an organised, elaborated and coherent response to personal questions concerning attachment figures.’ Humfress, H., O’Connor, T. G., Slaughter, J., Target, M., and Fonagy, P. (2002). ‘General and Relationship-Specific Models of Social Cognition: Explaining the Overlap and Discrepancies’. *Journal of Child Psychology and Psychiatry*, 43(7): 873–883, p. 880.

¹⁵ Choi-Kain, L. W., and Gunderson, J. G. (2008). ‘Mentalization: Ontogeny, Assessment, and Application in the Treatment of Borderline Personality Disorder’. *American Journal of Psychiatry*, 165(9): 1127–1135, p. 1133.

¹⁶ Target, M. (2008). ‘Commentary’, in F. N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (Psychoanalytic Inquiry Book Series: Volume 29), New York: Analytic Press, pp. 261–279. ‘Mentalisation is not a unitary capacity that a person either does or does not have. It is really a hierarchy of capacities’ (p. 270); Fonagy, P. and Luyten, P. (2009). ‘A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder’. *Development and Psychopathology*, 21(4): 1355–1381. ‘Research with this scale makes the further assumption that mentalization elicited in relation to past relationships will be related to current and even future relationships (Fonagy, Steele, Moran, Steele, & Higgitt, 1991). However, research has called into question the assumption that working models are traitlike. The substantial within-person variation in internal working models of others (e.g., father vs. mother) (e.g., Fraley, 2007; Pierce

the aggregate impression of mentalization and reflective function as unitary properties of a person. This aggregation, as it appeared in the context of the Adult Attachment Interview, was what was measured by the reflective functioning scale. But, in fact, substantial differences would be anticipated in the challenges individuals face in mentalizing, depending on their history and the characteristics of the present situation.

Fonagy, Gergely, and Target acknowledged that mentalization requires ‘a complex set of cognitive capacities’. These include the abilities: ‘1) to represent causal mental states of others with counterfactual contents (false beliefs), 2) to represent causal mental states of others with fictional contents (pretence, imagination, fantasy), 3) to simultaneously represent and differentiate between the mental models of the self and of the other about reality, 4) to infer and attribute the mental states of others from visible behavioural cues, and 5) to detect our own perceptible (behavioural, physiological, emotional, arousal, etc.) cues in order to infer, interpret, and attribute mental states to our self.’¹⁷ It by no means could be assumed that these cognitive capacities would have the same developmental trajectory, neurological supports, or social scaffolding for their elaboration. The conditions that would help one capacity flourish might even at times inhibit one or all of the other four. The distinctions between the five cognitive sub-capacities of mentalization suggest important individual differences in the capacity to mentalize the self compared with the capacity to mentalize others, and to distinguish these. However, they also imply that individuals may differ in the cognitive supports that permit the imaginative capacity to envision and represent counterfactual mental states, and the ability to detect observable cues for inferring and interpreting mental states.

A second major self-criticism raised by Fonagy and Target in *Psychoanalytic Theories* was the concern that, as a consequence of its origins in psychoanalysis, there ‘is over-emphasis on the earliest years as formative’ in their work to date.¹⁸ At this point, in 2003, Fonagy and colleagues still generally claimed that disorganized attachment, conceptualized as a kind of relational trauma, specifically would form the most important pathway to difficulties with mentalizing. When this trauma was severe and/or remained unresolved, the result could be anticipated to be BPD.¹⁹ However, this account did not sit well with the growing recognition of the variety of factors that could influence attentional control, affect regulation, and mentalization in general. In the mid-2000s, Fonagy and colleagues—Bateman especially—were struck by growing evidence from longitudinal research suggesting that the majority of patients with BPD do not have a history of sexual or physical abuse, though these experiences are certainly more common in this population.²⁰ Childhood and adolescence were clearly

& Lydon, 2001) supports a view of internal working models as hierarchically organized networks’ (p. 1374). For a recent re-affirmation that the mentalisation construct remains obscure, see Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., and Fonagy, P. (2020). ‘Learning to Mentalize: A Mediation Approach for Caregivers and Therapists’. *Clinical Psychology: Science and Practice*, 27(3): e12334.

¹⁷ Fonagy, P., Gergely, G., and Target, M. (2007). ‘The Parent–Infant Dyad and the Construction of the Subjective Self’. *Journal of Child Psychology and Psychiatry*, 48(3–4): 288–328, p. 290.

¹⁸ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 282. The self-criticism was reiterated in Fonagy, P. and Target, M. (2006). ‘The mentalization-focused approach to self pathology’. *Journal of Personality Disorders*, 20: 544–576: ‘we have placed too much emphasis on parents (particularly mothers)’ (p. 562).

¹⁹ Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). ‘The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence’. *Psychoanalytic Inquiry*, 23(3): 412–459.

²⁰ Bateman, A. W. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorders: Mentalization Based Treatment*, Oxford: Oxford University Press, pp. 27–37.

important periods for the development of mental health symptoms.²¹ However, the exact developmental processes in question required further examination. Whereas Fonagy and colleagues had previously emphasized the integral role of attachment in the development of BPD, this claim was increasingly tempered. A prominent criticism of the emphasis on attachment was offered by Gergely at a conference in 2005. Gergely argued that mentalization capabilities develop alongside, not out of, qualitative differences in attachment.²² He also proposed a developmental role for epistemic trust (see Chapter 7).

Four poles of mentalizing

Responding to problems with the construct of ‘mentalization’ and the question of developmental trajectory, Fonagy worked with Patrick Luyten in the late 2000s to address these concerns. Drawing on a review by Lieberman of the neuroscientific basis of forms of social cognition, in a paper published in 2009, Fonagy and Luyten described four ‘polarities’ of mentalization: internal and external; affective and cognitive; self and other; implicit and explicit.²³ They proposed these polarities as the most significant areas of potential individual differences in mentalizing. Following Lieberman, they also speculated that they would be subserved by different neurological structures. Since they were introduced, the four polarities have been absolutely central to guidance on matters of clinical technique offered by Fonagy and colleagues over the subsequent 10 years. They have regarded it as extremely helpful for clinicians to consider whether patients can engage flexibly in different forms of processing: internal and external; affective and cognitive; self and other; implicit and explicit. When a patient seems to be stuck on one ‘side’ of the polarity, Fonagy and colleagues have advised that clinicians should consider attempting a ‘contrary move’ to support mentalization on the other ‘side.’²⁴ This has been generally well received, and regarded as a helpfully specific and coherent piece of guidance. Yet, as well as pragmatic guidance for clinical technique, the four polarities were also presented as an account of the fundamental aspects of mentalization as a construct, and have subsequently been discussed in these terms.

A first distinction drawn by Fonagy and Luyten was between attention to the internal thoughts and feelings of a person, and attention to their behaviour. This distinction appeared to be an attempt to respond to their clinical impression that many patients with BPD can be vigilant and effective in observing external behavioural cues and have specific capabilities to notice the emotional states of others. Experimental researchers had also begun to document that patients with BPD do no worse, and sometimes actually better, than controls in facial emotion recognition and in theory of mind tasks.²⁵ However, the distinction between

²¹ Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). ‘Lifetime Prevalence and Age-Of-Onset Distributions of DSM-IV disorders in the National Comorbidity Survey Replication’. *Archives of General Psychiatry*, 62(6): 593–602.

²² The conference paper was later published as Gergely, G. and Unoka, Z. (2008), ‘Attachment, Affect-Regulation and Mentalization: The Developmental Origins of the Representational Affective Self’, in C. Sharp, P. Fonagy, and I. Goodyer (eds), *Social Cognition and Developmental Psychopathology Social Cognition and Developmental Psychopathology*, Oxford: Oxford University Press, pp. 303–340.

²³ Fonagy, P. and Luyten, P. (2009). ‘A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder’. *Development and Psychopathology*, 21(4): 1355–1381. See also Lieberman, M. D. (2007). ‘Social Cognitive Neuroscience: A Review of Core Processes’. *Annual Review of Psychology*, 58: 259–289.

²⁴ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 197.

²⁵ Domes, G., Czeschnek, D., Weidler, F., Berger, C., Fast, K., and Herpertz, S. C. (2008). ‘Recognition of Facial Affect in Borderline Personality Disorder’. *Journal of Personality Disorders*, 22: 135–147; Arntz, A., Bernstein, D.,

external and internal mentalization was quite a confused one. Previously, attention directed towards external behavioural cues rather than their internal mental states was not considered part of mentalizing: in fact, it had been considered *specifically* as a form of non-mentalizing (see Chapter 5). In Fonagy and Luyten's proposal, this non-mentalizing was termed 'external mentalizing'. It was contrasted with the identification and interpretation of mental states, which was now officially termed 'internal mentalizing', but generally just 'mentalizing'. There was evident room for muddle here, with 'mentalizing' being used to mean non-mentalizing.

Part of the problem was the expansive semantic domain covered by the term 'mentalization'. Where mentalising meant *the capacity to notice and respond to mental states in others*, then vigilance in observing external behavioural cues could be an asset. However, where mentalisation was taken to mean the more general *capacity to describe and interpret thoughts, feelings and intentions in oneself and others*, then a focus on external behaviour at the expense of consideration of the motivations and inner life of the other or the self was non-mentalizing. Later, Fonagy and Bateman would claim that 'mentalising is optimal when the dimensions—for example, emotion and cognition, or representation of self and other—are in balance and nonmentalizing modes are inactive.'²⁶ It is telling that the example of internalizing and externalizing mentalizing is not chosen to illustrate the ideal state of balance, because 'external mentalizing' veers towards a form of non-mentalizing. Detection and interpretation of observable behavioural cues is clearly a helpful contributor to mentalization, but only when it is in the service of envisioning and representing mental states. And this is not 'balance' between two poles.

In Fonagy and Luyten's 2009 paper, another distinction was drawn between the 'poles' of cognitive and affective mentalizing. This was, in fact, more a formalization of an existing distinction than a new addition to the theory. Since his earliest writings, Fonagy had tended to use the phrase 'thoughts and feelings' as a synonym for mental states, and this usage became more frequent in his writings with Target.²⁷ In their 2009 paper, Fonagy and Luyten drew a distinction between the capacity to identify and interpret thoughts, and the capacity to identify and interpret feelings. The former was termed 'cognitive mentalizing' and the latter was termed 'affective mentalizing'. This echoed developments in Baron-Cohen's work, which distinguished between systematizing and empathizing forms of social cognition.²⁸ Cognitive mentalizing and affective mentalizing appear in Fonagy and Luyten's 2009 paper to reflect the *target* of the activity—thoughts or feelings. It was not in the first instance a distinction regarding *how* the thought or feeling was recognized, which was generally assumed to be aligned and therefore left undistinguished. In principle, mental states could be understood through various means, among them formal-deductive reasoning and embodied affective resonance. The former might often be used to interpret thoughts. However,

Oorschot, M., and Schobre, P. (2009). 'Theory of Mind in Borderline and Cluster-C Personality Disorder'. *Journal of Nervous and Mental Disease*, 197: 801–807.

²⁶ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. viii.

²⁷ E.g. Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients'. *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 100; Target, M. and Fonagy, P. (1996). 'Playing with Reality II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479, p. 473; Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81: 853–873, p. 854.

²⁸ Baron-Cohen, S., Golan, O., Chakrabarti, B., and Belmonte, M. K. (2008). 'Social Cognition and Autism Spectrum Conditions', in C. Sharp, P. Fonagy, and I. Goodyer (eds), *Social Cognition and Developmental Psychopathology*, Oxford: Oxford University Press, pp. 29–56.

it can certainly be applied to feelings. In the latter case, then technically, according to Fonagy and Luyten's 2009 paper, formal-deductive reasoning would then be affective mentalizing. Likewise, embodied affective resonance is often used to interpret feelings, but might also be used to understand thoughts. In the latter case, technically embodied affective resonance would then be cognitive mentalizing. In sum, the characterization of cognitive and affective mentalizing by Fonagy and Luyten failed to draw a distinction between the cognitive/affective *means of understanding* and the *target of mentalization*. This has caused serious problems for subsequent work in discussions of phenomena like empathy, with a lack of clarity in the respective roles of cognition and affect, and the relative interdependence and independence of mentalizing and empathy.²⁹ Sharper attentiveness in the work of Jurist between affect as the means of understanding and as the target of mentalization has yet to filter effectively into wider discussions in the mentalization literature of the cognitive/affective 'poles'.³⁰

A third distinction drawn by Fonagy and Luyten was between mentalizing the self and mentalizing others. However, this was, by their own admission, not actually an opposition of the same kind as the others. When Fonagy and colleagues first announced the reflective functioning scale in 1991, they had acknowledged that—in contrast to Main's metacognitive monitoring scale—their reflective functioning scale collapsed differences between people who were good at reflecting on their own mental states, and those who were good at reflecting on the mental states of others.³¹ Yet, at the time, and through to the late 1990s, Fonagy and colleagues argued that mentalization of self and others was ultimately integrated into a unitary process in the course of development, which meant that it was reasonable to measure them all together with a unitary scale.³² In their 2009 paper, Fonagy and Luyten still held that the capacity to mentalize the self and others would generally come together. The reason for this was that they held that the capacity to understand the self is secondary, essentially an extension of the capacity to understand others in the context of attachment relationships. They held that it would be impossible to have a good understanding of the self and a poor understanding of others.³³ This claim appears to have been backtracked in

²⁹ Cf. Borelli, J. L., Stern, J. A., Marvin, M. J., Smiley, P. A., Pettit, C., and Samudio, M. (2020). 'Reflective Functioning and Empathy among Mothers of School-Aged Children: Charting the Space Between.' *Emotion*, Early View. As such, the relationship and potential overlap between mentalization-based therapy (MBT) and therapeutic modalities such as compassion-focused therapy remain to be worked out. This is a question currently pursued by Steve Pilling and collaborators: 'Compassion and mentalization: complementary, overlapping or opposed?'. Accessed at: <https://www.ucl.ac.uk/psychoanalysis/events/2021/may/compassion-and-mentalization-complementary-overlapping-or-opposed>.

³⁰ Jurist, E. L. (2005). 'Mentalized Affectivity'. *Psychoanalytic Psychology*, 22(3): 426–444.

³¹ Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). 'The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment'. *Infant Mental Health Journal*, 12(3): 201–218: 'The primary difference lies in the emphasis in Main's scale upon reflection by the subject on his or her own cognitive capacity (Flavell, 1976). In contrast, our scale emphasizes the use of psychological constructs to understand the behavior of self and other. The advantage of Main's scale is that it does not confound the subject's tendency to self-reflection with his or her accuracy in so doing. Our scale, by collapsing accuracy and habitual mode of thinking runs the risk of ascribing a high degree of reflective self function to individuals who are able to tell a plausible but inaccurate psychological story' (p. 215).

³² Fonagy, P. and Target, M. (1997). 'Attachment and Reflective Function: Their Role in Self-Organization'. *Development and Psychopathology*, 9(4): 679–700: 'The "fractionation" or splitting of all abilities as a function of tasks and domains is well demonstrated and we might expect reflective function to be subject to the same kind of developmental decalage (unevenness) which characterizes the rest of cognitive development ... Reflective function does not begin as a general capacity but is a particular skill tied to the task and domain where it is learned ... Normal development is from fractionation towards integration, which involves the coordination of previously separate skills' (pp. 695–696).

³³ For evaluation of Fonagy and Luyten's position on self–other interaction in the formation of subjectivity, see Liljenfors, R. and Lundh, L. G. (2015). 'Mentalization and Intersubjectivity towards a Theoretical Integration'. *Psychoanalytic Psychology*, 32(1): 36–60.

subsequent work, or even reversed into the proposition that “to mentalise others requires the capacity to mentalise the self.”³⁴ The developmental precedence and the causal precedence of mentalizing others and mentalizing the self has been an ongoing and unsolved problem in the work of Fonagy and colleagues. Nonetheless, there is clearly a distinction here. In the Fonagy and Luyten 2009 paper, as well as in subsequent work, the authors have urged attention to the possibility that individuals may have better understanding of others than of themselves, or poor understanding of both—for instance, through confusion of the mental states of self and other.

Perhaps the most important of the four distinctions drawn by Fonagy and Luyten was between automatic and controlled mentalizing, a distinction drawn from existing discussions in social cognition. There had been growing attention in the 2000s in cognitive neuroscience to the distinction between automatic and controlled mental processes.³⁵ Drawing on this literature, in 2004, Frith had argued that people at the high-functioning end of the autistic spectrum may lack automatic mentalizing, but have acquired skills in explicit mentalizing on the basis of experience and reasoning.³⁶ The distinction between automatic and controlled mentalizing appealed to Fonagy and Luyten’s mutual commitment to psychoanalysis, and the importance of non-conscious processes, including clinicians’ own implicit understandings.³⁷ The characterization of automatic mentalizing offered by Fonagy and Luyten has dissatisfied many commentators as insufficiently fleshed out.³⁸ However, in general terms, it would appear that they regarded automatic mentalizing as a quick, affect-led, unintended

³⁴ Bateman, A., Fonagy, P., and Campbell, C. (2019). ‘Antisocial Personality Disorder in Community and Prison Settings’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 335–349, p. 347. One way of interpreting this claim would be that Bateman and colleagues are here describing one part of a spiral. This is implied by Fonagy, P. and Allison, E. (2018). ‘The Origin of Human Life: A Psychoanalytic Developmental Perspective’, European Psychoanalytical Federation, 31st Annual Conference, Warsaw, 24 March: ‘Our model suggests that our understanding of ourselves emerges from an internalization of the understanding that others have of us (the primary caregiver). In order to understand them, however, we rely on self-understanding. If our self-understanding is good enough this creates the potential to establish a virtuous cycle. The better I understand myself, the better I will be able to understand my caregiver understanding me; this will enable me to elaborate my understanding of myself, in turn improving my capacity to grasp what she is teaching me about myself, and so on.’

³⁵ Satpute, A. B. and Lieberman, M. D. (2006). ‘Integrating Automatic and Controlled Processes into Neurocognitive Models of Social Cognition’. *Brain Research*, 1079: 86–97.

³⁶ Frith, U. (2004). ‘Emanuel Miller Lecture: Confusions and Controversies about Asperger Syndrome’. *Journal of Child Psychology and Psychiatry*, 45(4): 672–686.

³⁷ E.g. Fonagy, P. (2003). ‘Epilogue’. *Bulletin of the Menninger Clinic*, 67: 271–280: ‘As clinicians, our helpfulness to our patients may have more to do with the quality of implicit mentalization we offer through a general attitude rather than explicit elaboration of specific mental contents as has been frequently suggested in the past’ (p. 271). The line of influence is not clear, but this would seem well aligned with Bion’s recommendation that psychoanalysts should attempt to avoid constructing determinate thoughts and articulated feelings in sessions in order to keep themselves open for implicit understandings. Ogden, T. H. (2015). ‘Intuiting the Truth of What’s Happening: On Bion’s “Notes on Memory and Desire”’. *Psychoanalytic Quarterly*, 84(2): 285–306. See also Bion, W. R. (1970). *Attention and Interpretation*, London: Karnac Books: ‘There is the possibility of suppressing one or all of these functions of memory, desire, understanding, and sense either together or in turn. Practice in suppression of these faculties may lead to an ability to suppress one or other according to need, so that suspension of one might enhance the effect of domination by the other’ (p. 44).

³⁸ E.g. Køster, A. (2017). ‘Mentalization, Embodiment, and Narrative: Critical Comments on the Social Ontology of Mentalization Theory’. *Theory & Psychology*, 27(4): 458–476: ‘implicit mentalization suffers the somewhat unfortunate fate within MT of being attributed a gross significance in the literature while never receiving a systematic theoretical articulation’ (p. 464). Perhaps of particular importance for conceptualizing the distinction between the two, it remains unexamined whether the object of controlled mentalizing is always first an object of automatic mentalizing, as Hegel would suggest. This is hinted at in Shai, D. and Fonagy, P. (2014). ‘Beyond Words: Parental Embodied Mentalizing and the Parent Infant Dance’, in M. Mikulincer and P. R. Shaver (eds), *Mechanisms of Social Connections: From Brain to Group*, Washington, DC: American Psychological Association, pp. 185–203. However, questions of embodied mentalizing, even if acknowledged, have not been well integrated in subsequent work by Fonagy and colleagues.

process of identifying and making use of knowledge of mental states, running in the background of lived experience without awareness or effort, and responsive to the challenges of the particular moment. Controlled mentalizing seems to have been regarded as the deliberate, cognition-led and effortful attempt to identify and reason about mental states, a relatively slow and serial process, somewhat more removed from the hurly-burly of the demands of the moment.

The way that the distinction between automatic and controlled processing was introduced by Fonagy and Luyten gives the impression that automatic processing is the same as implicit, non-mentalizing processing. However, elsewhere in the article, it is implied that it is possible to have i) automatic mentalizing and ii) automatic non-mentalizing; just as it is possible to have iii) controlled mentalizing and iv) controlled non-mentalizing. Havsteen-Franklin has recently drawn out these distinctions, and termed this the ‘mentalizing quadrant’.³⁹ It is very confusing for Fonagy and Luyten to refer to both automatic mentalizing and automatic non-mentalizing as ‘automatic mentalizing’. It forces them into strange claims, such as that the fight/flight/freeze response is a form of mentalizing.⁴⁰ Or again, consider that Fonagy and colleagues have described the reflective functioning scale for the Adult Attachment Interview as measuring automatic mentalizing.⁴¹ This would simply be contradictory if by this they meant implicit, non-mentalizing processing.⁴² In brief, Fonagy and Luyten’s primary characterization of automatic and controlled mentalization presumed alignment with non-mentalizing and mentalizing, when even statements in their own article suggest a more complicated situation.⁴³ It also presumed alignment of automatic with fast, and controlled with slow, mentalizing. This provides no place for the implicit, slow processes of ‘feeling things out’ that precisely, as Fonagy has acknowledged elsewhere, comprise a large bulk of work in psychotherapy.⁴⁴ It also leaves unclear the respective role of culture in

³⁹ Havsteen-Franklin, D. (2019). ‘Creative Arts Therapies’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice*, pp. 181–195, Washington, DC: American Psychiatric Association, Figure 11.1. This point is also suggested in Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment For Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 7–9. See also Davidsen, A. S. and Fosgerau, C. F. (2015). ‘Grasping the Process of Implicit Mentalization’. *Theory & Psychology*, 25(4): 434–454.

⁴⁰ Luyten, P., Malcorps, S., Fonagy, P. and Ensink, K. (2019). ‘Mentalising and Trauma’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 79–102, p. 84.

⁴¹ E.g. Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment For Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 111.

⁴² Allen has likewise characterized implicit mentalizing in terms that suggest that it *is* mentalising, and distinct from implicit non-mentalizing: ‘We are mentalizing implicitly when we empathize intuitively and nonverbally, “mirroring” others’ emotional states. We mentalize implicitly when we respond with a look of interest to what our patient just said, perhaps leaning forward a bit and raising our brow.’ Allen, J. G. (2003). ‘Mentalizing’. *Bulletin of the Menninger Clinic*, 67(2): 91–112, p. 95.

⁴³ Recently, Fonagy and colleagues have tended to refer less to automatic and controlled mentalizing, and instead to ‘fast’ and ‘slow’ mentalising. This follows Kahneman, D. (2011). *Thinking, Fast and Slow*, London: Macmillan. Fonagy, P., Allison, E., and Campbell, C. (2019). ‘Mentalising, Resilience and Epistemic Trust’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 68.

⁴⁴ Fonagy, P. (2003). ‘Epilogue’. *Bulletin of the Menninger Clinic*, 67(3): 271–280. See also Jurist, E. L. (2005). ‘Mentalized Affectivity’. *Psychoanalytic Psychology*, 22(3): 426–444. Recently it has been acknowledged that ‘MBT is constrained by its representational conceptualization’. Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., and Fonagy, P. (2020). ‘Learning to Mentalize: A Mediational Approach for Caregivers and Therapists’. *Clinical Psychology: Science and Practice*, 27(3): e12334. The focus of MBT on representational processes has various roots, among them the development of ideas of mentalizing from work using the Adult Attachment Interview (see Chapter 3). However, an additional factor is the lack of space in the theory for considering implicit, slow processes of ‘feeling things out’, which occur without explicit representational content and sometimes without a discrete intention.

implicit mentalizing—perhaps as conventional response, perhaps as habit—and controlled mentalizing—perhaps as semiotic scaffolding for reflection. This makes automatic and controlled mentalization appear to be solely kinds of individual response, cutting them out from their constitutive contexts. As we shall see in Chapter 9, Fonagy and colleagues have spotlighted such processes as central to their current thinking about mentalization; however, they are yet to reconsider automatic and controlled mentalizing on this basis.

Fonagy has spoken of his dismay at the lack of uptake of the 2009 model of four dimensions by other researchers and clinicians, and he has speculated that the model may be too complicated.⁴⁵ We hold that, rather than too complicated, the lack of uptake results from the fact that the account is confusing on several grounds. For instance, the term ‘mentalizing’ in the account of automatic and controlled processes sometimes was used, precisely, to mean ‘non-mentalizing’. Likewise, ‘external mentalizing’ sometimes meant, precisely, ‘non-mentalizing’. The opposition between affective and cognitive mentalizing was also not especially sharp, because it was not clear whether what is ‘affective’ or ‘cognitive’ refers to the target of understanding or the mechanism for understanding. The opposition between automatic and controlled mentalizing also gives no clear place to the slow process of ‘feeling things out’, which is not well characterized either as automatic or controlled processing.

Leaving the four poles of mentalizing aside, the Fonagy and Luyten paper also offered an important revision of the developmental model. They argued that attachment relationships calibrate: i) the threshold at which an individual will feel tugged towards automatic processing under conditions of high arousal, ii) the extent to which automatic processing is able to identify and interpret mental states; iii) the ability of the individual to resist the tug of automatic processing and retain the capacity for explicit reasoning; and iv) the extent to which familiar others are sought in the context of high arousal.⁴⁶ Fonagy and Luyten anticipated that secure attachment would raise the threshold for activation of the attachment system, increase capacities for automatic mentalizing, stabilize the retention of explicit reasoning, and draw effectively on social support for the maintenance of emotion regulation. Just like the secure infant in the Strange Situation can shift readily between exploration and care-seeking, depending on the context, Fonagy and Luyten presumed that secure attachment would facilitate the easy shift of gears between automatic and controlled mentalizing depending on what was called for.⁴⁷

Avoidant attachment would raise the threshold for activation of the attachment system, reduce capacities for automatic mentalizing about the self, stabilize the retention of explicit reasoning, and make less use of familiar others in achieving emotion regulation. However, this explicit reasoning could still have gaps or lapses in mentalizing where it would otherwise

⁴⁵ Duschinsky, R., Collver, J., and Carel, H. (2019). “‘Trust Comes from a Sense of Feeling One’s Self Understood by Another Mind’: An Interview with Peter Fonagy”. *Psychoanalytic Psychology*, 36(3): 224–227.

⁴⁶ Fonagy, P. and Luyten, P. (2009). ‘A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder’. *Development and Psychopathology*, 21(4): 1355–1381: ‘Neuroimaging studies support the notion that attachment history affects the setting of the “switch,” which turns the mentalizing system from planned, controlled and organized cognition to automatic processing with narrowed, poorly sustained attention, and increased vigilance for attachment disruptions such as rejection and abandonment’ (p. 1368).

⁴⁷ The point is made again, and somewhat more clearly, in Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: ‘Particularly in a secure attachment environment, when things are running smoothly on an interpersonal level, more deliberate or controlled mentalising is not called for; in fact, the use of such a mentalising style might hinder such interactions, making them feel unduly weighty or uncomfortably overwrought (hypermentalised)’ (p. 8); ‘Well functioning mentalising involves the ability to switch flexibly and responsively from automatic to controlled mentalising’ (p. 9).

need to be informed by affect and automatic mentalizing.⁴⁸ By contrast, ambivalent/resistant attachment and disorganized attachment would be anticipated to i) lower the threshold for activation of the attachment system, ii) reduce capacities for automatic mentalizing about others and prompt self/other confusions, iii) destabilize the retention of explicit reasoning, and iv) contribute to difficulties in making use of familiar others for emotion regulation.⁴⁹ This configuration is of particular importance for Fonagy and Luyten, because they regarded these interacting elements specifically as the basis for the development of BPD. There may be few deficits in mentalizing for patients with BPD when they are calm. But it was anticipated—and indeed later research would find—that in response to arousal, and specifically the activation of the attachment system, the anticipated difficulties in mentalizing would become evident.⁵⁰ Early disruptions in attachment would make mental states difficult to identify and attribute to self or other, while simultaneously prompting vigilance regarding mental states, which would be interpreted as potentially threatening. The lowered threshold for activation of the attachment system would keep controlled mentalization offline, hindering the integration of affective and cognitive forms of mentalizing. In turn, this would sustain difficulties in mentalizing the self and others, and ensuing difficulties in making use of social relationships for modulating troubling affects (see Figure 4.1).⁵¹

In short, Fonagy and Luyten's model suggested that difficulties with mentalising, affect regulation, social relationships and executive function can exacerbate one another. When arousal is low and there is little perception of threat few, if any, deficits in mentalisation are likely to be evident. However, when the individual faces difficult feelings or situations, Fonagy and Luyten proposed that the result will be a lowered threshold for entering a state of automatic non-mentalized responding, and a raised threshold for engaging in the controlled mentalized processing that might help integrate affective arousal with cognitive understanding and perspective. Automatic, non-mentalized responses are likely to be headlong, lacking in measure and composure, and predisposed to contribute to relationship difficulties, especially in attachment relationships.⁵² In this state, there is little patience for prompts

⁴⁸ See also Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'Explicit reflection cannot feel real unless it is contextualised by intuitive awareness of the mental states being reflected on' (p. 10).

⁴⁹ The distinction between ambivalent/resistant and disorganized attachment is not sharply drawn in the work of Fonagy and colleagues. This aligns with the work of Roisman and colleagues, whose psychometric studies have suggested that the two categories should be treated as contributory to a single dimension of dysregulated attachment. See e.g. Raby, K. L., Labella, M. H., Martin, J., Carlson, E. A. and Roisman, G. I. (2017). 'Childhood Abuse and Neglect and Insecure Attachment States of Mind in Adulthood: Prospective, Longitudinal Evidence from a High-Risk Sample'. *Development and Psychopathology*, 29(2): 347–363. Other attachment researchers have contested this position e.g. van IJzendoorn, M. H. and Bakermans-Kranenburg, M. J. (2014). 'Confined Quest for Continuity: The Categorical versus Continuous Nature of Attachment'. *Monographs of the Society for Research in Child Development*, 79(3): 157–167. It is interesting, in this regard, that though borderline individuals are over-represented in the preoccupied classification of the Adult Attachment Interview on a three-way analysis, when the unresolved classification is included, there is no over-representation in the preoccupied classification—only in the unresolved classification. This finding could in principle be used to support either position. Bakermans-Kranenburg, M. J. and van IJzendoorn, M. H. (2009). 'The First 10,000 Adult Attachment Interviews: Distributions of Adult Attachment Representations in Clinical and Non-Clinical Groups'. *Attachment & Human Development*, 11(3): 223–263, p. 230.

⁵⁰ E.g. Dixon-Gordon, K. L., Chapman, A. L., Lovasz, N., and Walters, K. (2011). 'Too Upset to Think: The Interplay of Borderline Personality Features, Negative Emotions, and Social Problem Solving in the Laboratory'. *Personality Disorders: Theory, Research, and Treatment*, 2: 243–260; Sharp, C. (2014). 'The Social-Cognitive Basis of BPD: A Theory of Hypermentalizing', in C. Sharp and J. L. Tackett (eds), *Handbook of Borderline Personality Disorder in Children and Adolescents, Volume 5*, New York: Springer, pp. 211–225.

⁵¹ Fonagy, P. and Luyten, P. (2009). 'A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder'. *Development and Psychopathology*, 21(4): 1355–1381.

⁵² See also Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre: 'there is no context more likely to induce a loss of mentalizing than

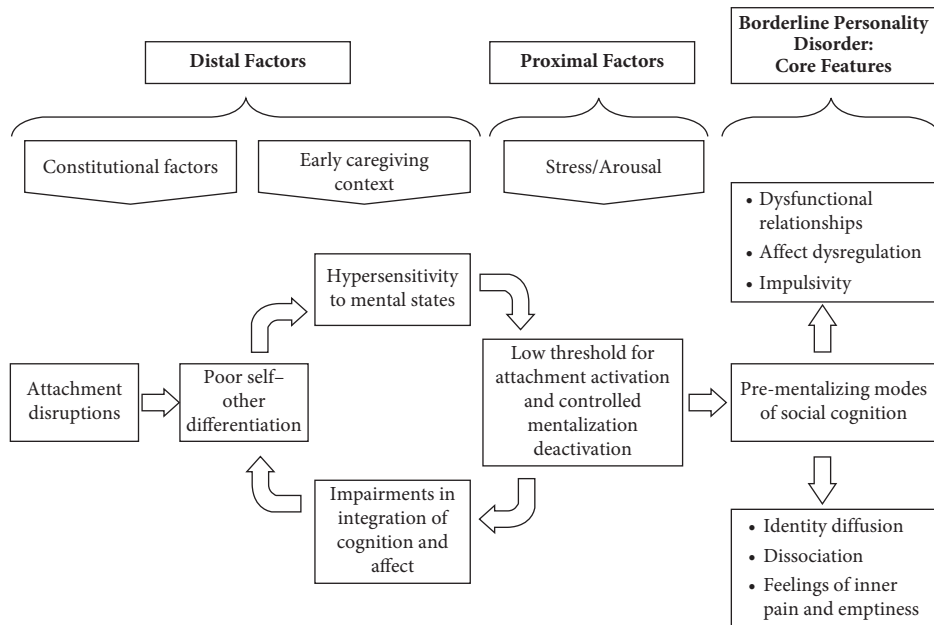


Figure 4.1 Model of the causes and features of BPD.

Source: Reproduced from Peter Fonagy and Patrick Luyten, 'A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder', *Development and Psychopathology* 21(4): 1355–1381, Figure 2, DOI: <https://doi.org/10.1017/S0954579409990198> Copyright © Cambridge University Press 2009.

to stop and reconsider before acting. Repeated or sustained deployment of these responses will also hinder the capacity of the individual to identify and understand their own mental states and experience fulfilment from their relationships, contributing to identity diffusion and feelings of inner emptiness.

In 2012, Luyten, Fonagy, and colleagues built on this account to propose a multidimensional approach to the assessment of mentalizing.⁵³ They advised that psychological assessment should appraise the extent to which the mentalizing profile differs across specific attachment relationships. Some relationships may prompt lowered or raised thresholds for states of automatic, non-mentalized processing, and may hinder or support the maintenance of the capacity for controlled, mentalized processing. They emphasized the importance of the extent to which an individual can accept help and support in sustaining mentalizing under conditions of perceived threat or arousal. They also presented the concept of a 'mentalizing profile' as a characterization of how an individual functions on each of the four axes: whether the individual attends only to external behavioural cues, or also considers internal feelings, thoughts, and motivations; whether the individual can describe and interpret more readily thoughts or feelings; whether the individual has greater ability to consider mental states in

attachment relationships of all kinds, including family interactions. It is here that relationships tend to be at their most fraught, their most loving and their most intense emotionally, and so the stage is set on a daily basis for interactions that potentially stimulate a loss of mentalizing in one or more members of the family.'

⁵³ Luyten, P., Fonagy, P., Lowyck, B., and Vermote, R. (2012). 'The Assessment of Mentalization', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 43–66.

others or also to understand mental states in themselves; and how readily the individual enters into states of automatic, non-mentalized processing, as well as whether they can readily access states of controlled, mentalized processing. However, without any measure available of the four poles of mentalizing, the concept of the mentalizing profile has remained a recommendation for clinical practice based much more in theory than in dedicated empirical research. This has made claims about the mentalizing profiles of different client groups, or the treatments that should be tailored to different mentalizing profiles, essentially unfalsifiable. Given the problems with the four poles framework, it is not certain that creation of such a measure is even technically possible without significant further revision to the framework. Nonetheless, the proposal to attend to mentalizing profiles represents an advance in advocating a reduction of the reification of mentalizing in clinical practice.

In line with this concern for greater specificity in conceptualizing mentalization, Berthelot, Fonagy, and colleagues attempted a replication of the Fonagy, Steele, Moran, Steele, and Higgitt study of intergenerational transmission, with a sample of 57 mothers screened for high probability of inadequate care in their own childhoods.⁵⁴ Adult Attachment Interviews and Strange Situation procedures were conducted, as in the original study. However, as well as using the reflective functioning scale on the Adult Attachment Interviews conducted with the mothers, the researchers also elaborated a sub-scale focused specifically on the capacity of the speaker to mentalize about traumatic episodes at any point in their attachment history. As in the original study, in 72% of cases, there was a match between the Adult Attachment Interview classification of the caregiver and Strange Situation classifications of the infant–caregiver dyad on security/insecurity. In the case of speakers unresolved for loss or trauma, 70% had relationships classified as disorganized in the Strange Situation procedure. In a regression, unresolved trauma accounted for 22% of variance in infant attachment disorganization. The general reflective function scale made no additional contribution to predicted variance. Yet, when reflective function about attachment-related trauma was entered into the regression, the model accounted for 41% of variance. The researchers concluded that it was not lack of mentalization in general that was influential for predisposing disorganized infant–caregiver attachment relationships, but the caregiver's capacity to mentalize about specific attachment-related traumas.⁵⁵

Another qualification to the Fonagy and Luyten model came through Fonagy's collaboration with Karin Ensink and colleagues. Fonagy and Luyten had implied an undifferentiated contribution of non-mentalizing to mental health symptoms. Yet the pathway to impulsive and aggressive behaviour might not be the same as to inner emptiness and depression. This is hinted at by Figure 2 in Fonagy and Luyten's paper, where internalizing and externalizing symptoms were placed in different boxes. But the distinction was not drawn in the text.⁵⁶ Ensink and colleagues conducted a study with 74 mother–child dyads where the child had experienced sexual abuse, and 96 matched controls.⁵⁷ The children in the study were aged

⁵⁴ Berthelot, N., Ensink, K., Bernazzani, O., Normandin, L., Luyten, P., and Fonagy, P. (2015). 'Intergenerational Transmission of Attachment in Abused and Neglected Mothers: The Role of Trauma-Specific Reflective Functioning'. *Infant Mental Health Journal*, 36(2): 200–212.

⁵⁵ See also Borelli, J. L., Cohen, C., Pettit, C., Normandin, L., Target, M., Fonagy, P., and Ensink, K. (2019). 'Maternal and Child Sexual Abuse History: An Intergenerational Exploration of Children's Adjustment and Maternal Trauma-Reflective Functioning'. *Frontiers in Psychology*, 10.

⁵⁶ The researchers have recently conducted factor analytic work on self-report items of depression, identifying a distinct factor in which depression is accompanied by aggression. Rost, F., Luyten, P., and Fonagy, P. (2018). 'The Anaclitic–Introjective Depression Assessment: Development and Preliminary Validity of an Observer-Rated Measure'. *Clinical Psychology & Psychotherapy*, 25(2): 195–209.

⁵⁷ Ensink, K., Bégin, M., Normandin, N., Godbout, N., and Fonagy, P. (2017). 'Mentalization and Dissociation in the Context of Trauma: Implications for Child Psychopathology'. *Journal of Trauma & Dissociation*, 18(1): 11–30.

7–12 and completed the adaptation of the Adult Attachment Interview for this age group, which was coded for age-appropriate reflective function. Child dissociative, externalizing, and inappropriate sexual behaviours were assessed by parent report; depression was measured by child self-report. Path analysis modelling revealed that the association between sexual abuse and impulsivity and aggression accounted for around 60% of variance, and was entirely mediated by child reflective function and dissociation. Most of the mediation was accounted for by dissociation. However, there was also some sequential mediation through the effects of sexual abuse on reflective function, which in turn contributed to dissociation. Similar findings were found for inappropriate sexual behaviours. By contrast, 34% of the effects of abuse on child depression could be accounted for by child reflective functioning, and here dissociation had no mediating role. Ensink and colleagues concluded from their findings that the Fonagy and Luyten model needed to be rectified to specify the role of dissociation rather than merely absorbing it into non-mentalizing, given the finding of serial mediation between reflective function and dissociation in predicting impulsivity and aggression. Fonagy has admitted that dissociation has tended to be a rather ‘shadowy concept’, and requires renewed discrimination and attention from non-mentalizing.⁵⁸ This has been facilitated in recent years by the availability of a self-report measure of mentalizing (see Chapter 5).⁵⁹

A response to the limitations of the Fonagy and Luyten 2009 account was presented by Fonagy, Luyten, Allison, and Campbell in 2017 in a landmark paper entitled ‘What We have Changed our Minds about.’⁶⁰ The major amendment made there to the model of the genesis of BPD was to qualify the importance of early attachment. Instead of insisting on the importance of early disorganized attachment (see Chapter 3), Fonagy and colleagues argued for a cycle that could begin at any point between problems in relationships, difficulties in sustaining attention to and understanding mental states, and heightened negative affect. Problems in attachment relationships across childhood and adolescence might contribute to all three processes, but without attachment being regarded as a foundational cause, as in the Luyten and Fonagy 2009 model. It was now proposed that the roots of borderline as a ‘personality’ disorder lay not in the long-term influence of attachment or in the stability of personality, but in the stability of the interaction between problems in relationships, difficulties in sustaining attention to and understanding mental states, and heightened negative affect. This account will be discussed further in Chapter 7.⁶¹ Another shift from 2009 to 2017 has been to stress the importance of social interaction in stabilizing or destabilizing mentalizing capacities. In particular, Fonagy’s collaboration with Liz Allison, Chloe Campbell, and Patrick Luyten had led to a perspective that does not treat mentalization as a good in itself,

⁵⁸ Fonagy, P. (2018). ‘Preface’ to Werner Bohleber, *Destructiveness, Intersubjectivity and Trauma: The Identity Crisis of Modern Psychoanalysis*, London: Routledge, pp. xi–xiv, p. xii. See also Schimmenti, A. (2015). ‘Behind the Closed Doors of Mentalizing. A Commentary on “Another Step Closer to Measuring the Ghosts in the Nursery: Preliminary Validation of the Trauma Reflective Functioning Scale”’. *Frontiers in Psychology*, 6: 380; Ensink, K., Fonagy, P., Berthelot, N., Normandin, L., and Bernazzani, O. (2015). ‘Response: Behind the Closed Doors of Mentalizing. A Commentary on “Another Step Closer to Measuring the Ghosts in the Nursery: Preliminary Validation of the Trauma Reflective Functioning Scale”’. *Frontiers in Psychology*, 6: 697.

⁵⁹ E.g. Huang, Y. L., Fonagy, P., Feigenbaum, J., Montague, P. R., Nolte, T., and Mood Disorder Research Consortium (2020). ‘Multidirectional Pathways between Attachment, Mentalizing, and Posttraumatic Stress Symptomatology in the Context of Childhood Trauma’. *Psychopathology*, 53: 48–58.

⁶⁰ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). ‘What We have Changed our Minds about. Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication’. *Borderline Personality Disorder and Emotion Dysregulation*, 4: 9.

⁶¹ See also Luyten, P., Campbell, C., and Fonagy, P. (2020). ‘Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach’. *Journal of Personality*, 88(1): 88–105.

but as a catalyst to allow an individual to sift and learn from social experience, effectively and with good discrimination. This will be discussed further in Chapter 9.

Defining mentalizing

In a recent commentary, Fonagy and colleagues recalled that ‘initially we provided a model proposing improvements in the capacity to mentalize as a central mechanism shared by a wide range of treatment approaches. Further reflection and empirical and clinical evidence have led us to doubt the sufficiency of this assumption. We have all seen many patients whose capacity to mentalize benefited from psychoanalytic psychotherapy without corresponding improvement in their social-emotional functioning. It seems, then, that we should be more specific.’⁶² Even allies of Fonagy and colleagues have come to characterize the concept of ‘mentalization’ as functionally ‘bloodless’ now.⁶³ Such a characterization takes matters too far. The concept remains very much alive and purposive. However, to support future attempts at greater specificity, as well as effective articulation and review of ideas in the rest of the book, to close this chapter we will present a survey of definitions of mentalization over 30 years: from Fonagy’s first use of the term in 1989 to the present. Our discussion of the different definitions is certainly not an end in itself; instead, our intention is to be constructive in attempting to map the scope of the concept, on the way to a new synthesis and definition. This will then support an analysis of the central elements of the construct, to facilitate the discussion of the subsequent chapters. Twenty-eight distinct definitions of mentalization can be identified in the writings of Fonagy and his collaborators. These can be seen in Table 4.1.

Remarks on the definitions of mentalizing:

i. Conceiving and reconsidering: The initial definition, in 1989, of the capacity to ‘conceive’ of mental states, was too narrow. The capacity to conceive of mental states would not be helpful to an individual if these conceptions could not then be employed. A second attempt can be seen in the definition, in 1995, of mentalization as the capacity to think about mental states, which also presumably contains the capacity to conceive of the states thought about. From 1999, this is further revised to the capacity to ‘understand and interpret’ the mental states. This is then relatively stable over time. The 2012 definition unpacks interpreting mental states as entailing two processes: ‘describing’ and ‘explaining’. This is different from behaviour, which is to be ‘understood’ in terms of mental states. Understanding, then, appears to be the successful outcome of describing and explaining behaviour in terms of mental states, though this is not pinned down. The definitions do not imply that affectively attuning to mental states is construed as mentalizing, except insofar as it is encompassed by conceiving of or reconsidering mental states.

ii. No future: From the late 1990s, definitions stress that thinking, understanding, or interpreting are all integrative forms of mental processing. From the mid-2000s, we see a specification of integration of experiences and representation as a requisite component of mentalization. The nature of representation here is not specified, but would seem entailed by thinking, understanding, and interpreting. Across the definitions, it seems to be implied that the perceptual experiences that mentalization targets may occur in the present or be drawn

⁶² Fonagy, P., Allison, E. and Campbell, C. (2019). ‘Commentary on ““Trust Comes from a Sense of Feeling One’s Self Understood by Another Mind”: An Interview with Peter Fonagy”’. *Psychoanalytic Psychology*, 36(3): 228.

⁶³ See e.g. Auerbach, J. Allen, J. G., Fonagy, P., and Bateman, A. W. (2007) ‘Mentalizing in Clinical Practice (Book Review)’. Accessed at: <https://www.apadivisions.org/division-39/publications/reviews/mentalizing-2>.

Table 4.1 Definitions of mentalization

Year	Source	Definition of mentalization
1989	Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients'. <i>Bulletin of the Anna Freud Centre</i> , 12: 91–115, p. 97.	'For the sake of brevity rather than reification I would like to label the capacity to conceive of mental states in oneself, as the capacity to mentalize.'
1995	Fonagy, P. and Target, M. (1995). 'Understanding the Violent Patient: The Use of the Body and the Role of the Father'. <i>The International Journal of Psychoanalysis</i> , 76: 487–501, p. 487.	'Mentalisation, as the capacity to think about mental states'
1995	Fonagy, P. and Target (1995). 'Playing with Reality: The Development of Psychic Reality and its Malfunction in Borderline Personalities'. <i>The International Journal of Psychoanalysis</i> , 76: 39–44.	'With the capacity to mentalize, to see ideas as merely ideas and not facts, to play with different points of view, there comes the capacity to test ideas against reality and therefore to moderate their impact' (p. 229).
1996	Fonagy, P. (1996). 'The Significance of the Development of Metacognitive Control over Mental Representations in Parenting and Infant Development'. <i>Journal of Clinical Psychoanalysis</i> , 5(1): 67–86, p. 74.	'The psychological processes underpinning the view of oneself and others as motivated by mental states'
1997	Fonagy, P. (1997). 'Where Cure was Inconceivable. The Aims of Modern Psycho-Analysis with Borderline Patients'. <i>Texte</i> , 3(17): 11–25, p. 15.	'These representations were unmetabolized, in Bion's terms, non-mentalising within our theoretical framework.'
1997	Fonagy, P. ([1997] 2002). 'Multiple voices versus meta-cognition: An attachment theory perspective', in V. Sinason (ed.), <i>Attachment, Trauma and Multiplicity: Working with Dissociative Identity Disorder</i> London: Brunner-Routledge, pp. 71–85, p. 79.	'The symptom of dissociation itself, the predominance of multiple voices in treatment may be usefully seen as the converse of mentalization.'
1998	Fonagy, P. (1998). 'Prevention, the Appropriate Target of Infant Psychotherapy'. <i>Infant Mental Health Journal</i> , 19(2): 124–150, p. 136.	'Mentalizing, conceiving of interpersonal experience in terms of mental states or minds' (p. 136).
1999	Fonagy, P. (1999). 'Points of Contact and Divergence between Psychoanalytic and Attachment Theories: Is Psychoanalytic Theory Truly Different?' <i>Psychoanalytic Inquiry</i> , 19(4): 448–480, p. 461.	'The notion of reflective function or mentalization is already present in Freud's (1911) notion of Bindung, or linking. Bindung refers to the qualitative change from the physical (immediate) to the psychological (associative) quality of linking'

(continued)

Table 4.1 Continued

Year	Source	Definition of mentalization
1999	Fonagy, P. (1999). 'Male Perpetrators of Violence against Women: An Attachment Theory Perspective'. <i>Journal of Applied Psychoanalytic Studies</i> , 1(1): 7–27, p. 13.	'Mentalization, the capacity to understand and interpret human behavior in terms of the putative mental states underpinning it, arises through the experience of having been so understood in the context of an attachment relationship.'
2000	Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. <i>The International Journal of Psychoanalysis</i> , 81: 853–873, p. 854.	'The capacity to mentalize: to assume the existence of thoughts and feelings in oneself and in others, and to recognize these as connected to outer reality'
2002	Fonagy, P., Gergely, G., Jurist, E.L. and Target, M. (2002). <i>Affect Regulation, Mentalisation and the Development of the Self</i> , London: Karnac Books, p. 3.	'Mentalisation . . . is the process by which we realise that having a mind mediates our experience of the world'
2003	Allen, J. G. (2003). 'Mentalizing'. <i>Bulletin of the Menninger Clinic</i> , 67(2): 91–112, p. 94.	'Mentalizing entails interpreting the behavior of oneself and others in terms of intentional mental states, such as desires, feelings, beliefs, and the like'
2004	Bateman, A. W. and Fonagy, P. (2004). <i>Psychotherapy for Borderline Personality Disorders: Mentalization Based Treatment</i> , Oxford: Oxford University Press, p. 70.	'The capacity to think about mental states as separate from, yet potentially causing actions.'
2006	Fonagy, P. and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology'. <i>Journal of Personality Disorders</i> , 20: 544–576, p. 545.	'Mentalization involves both a self-reflective and an interpersonal component. It is underpinned by a large number of specific cognitive skills, including an understanding of emotional states, attention and effortful control, and the capacity to make judgements about subjective states as well as thinking explicitly about states of mind—what we might call mentalization proper. In combination, these functions enable the child to distinguish inner from outer reality and internal mental and emotional processes from interpersonal events.'
2007	Fonagy, P., Gergely, G., and Target, M. (2007). 'The Parent–Infant Dyad and the Construction of the Subjective Self'. <i>Journal of Child Psychology and Psychiatry</i> , 48(3–4): 288–328, p. 288.	'We define mentalization following a tradition in philosophy of mind established by Brentano (1973/1874), Dennett (1978) and others as a form of mostly preconscious imaginative mental activity, namely, perceiving and interpreting human behaviour in terms of intentional mental states (e.g., needs, desires, feelings, beliefs, goals, and reasons).'
2008	Bateman, A. and Fonagy, P. (2008). 'Comorbid Antisocial and Borderline Personality Disorders: Mentalization-Based Treatment'. <i>Journal of Clinical Psychology</i> , 64(2): 181–194, p. 182.	'Mentalizing simply implies a focus on mental states in oneself or in others, particularly in explanations of behaviour'

Table 4.1 Continued

Year	Source	Definition of mentalization
2008	Sharp, C. and Fonagy, P. (2008). 'The Parent's Capacity to Treat the Child as a Psychological Agent: Constructs, Measures and Implications for Developmental Psychopathology'. <i>Social Development</i> , 17(3): 737–754, p. 740.	'What exactly does mentalizing mean within this context? Slade (2005), writing from a psychoanalytic perspective, summarizes the concept as referring to both a cognitive process, akin to psychological insight or perspective-taking, and an emotional process, that is, the capacity to hold, regulate, and fully experience one's own and others' emotions in a non-defensive way without becoming overwhelmed or shutting down.'
2008	Bouchard, M. A., Target, M., Lecours, S., Fonagy, P., Tremblay, L. M., Schachter, A., and Stein, H. (2008). 'Mentalization in Adult Attachment Narratives: Reflective Functioning, Mental States, and Affect Elaboration Compared'. <i>Psychoanalytic Psychology</i> , 25(1): 47–66, p. 48.	'Mentalization is a term used by both developmental psychologists and psychoanalysts to refer to a core process of human social functioning and self-regulation, involved in the establishment of robust links between personally meaningful early experiences and their representation ... However, mentalization has been diversely construed and measured.'
2009	Slade, A. (2009). 'Mentalizing the Unmentalizable: Parenting Children on the Spectrum'. <i>Journal of Infant, Child, and Adolescent Psychotherapy</i> , 8(1): 7–21, p. 8.	'The term mentalization (or reflective functioning) refers to the capacity to envision mental states in the self or other, to use an understanding of mental states—intentions, feelings, thoughts, desires, and beliefs—to make sense of, and even more important, to anticipate another's (or her own) actions.'
2010	Fonagy, P. (2010). 'Attachment Trauma and Psychoanalysis: Where Psychoanalysis meets Neuroscience', in M. Leuzinger-Bohleber, J. Canestri, and M. Target (eds), <i>Early Development and its Disturbances: Clinical, Conceptual and Empirical Research on ADHD and other Psychopathologies and its Epistemological Reflections</i> , London: Karnac Books, pp. 53–75, p. 55.	'Mentalization is a concept originally introduced by French psychoanalysts (Luquet 1981, 1987; Marty and De M'Uzan 1963) working with psychosomatic patients. They noted a lack of symbolization of mental states in such individuals, a lack of freedom in free association, and a characteristic way of thinking too close to sensations and primary unconscious fantasies.'
2012	Asen, E. and Fonagy, P. (2012). 'Mentalization-Based Therapeutic Interventions for Families'. <i>Journal of Family Therapy</i> , 34(4): 347–370, p. 347.	'Mentalizing is a process and it generally occurs without effort or specific consciousness. It can be summarized as seeing ourselves from the outside and seeing others from the inside.'
2012	Fonagy, P., Bateman, A. W. and Luyten, P. (2012). 'Introduction and Overview', in <i>Handbook of Mentalizing in Mental Health Practice</i> . Washington, DC: American Psychiatric Publishing, pp. 3–42, p. 3.	'Mentalising is a term that has been introduced simultaneously into neuroscience and clinical thinking to denote the remarkable and pervasive human tendency to look beyond the visible shell of the body in understanding behaviour and seeking descriptions and explanations in terms of states of mind.'

(continued)

Table 4.1 Continued

Year	Source	Definition of mentalization
2015	Target, M. (2015). 'A Developmental Model of Sexual Excitement, Desire and Alienation.' <i>Sexualities: Contemporary Psychoanalytic Perspectives</i> , Alessandra Lemma and Paul E. Lynch (eds), London: Karnac Books, pp. 43–62, p. 46.	'Mentalization is the conscious or preconscious recognition that behaviour is understandable given underlying mental states and intentions, that it therefore has motivation and meaning.'
2016	Bateman, A. W. and Fonagy, P. (2016). <i>Mentalization-Based Treatment for Personality Disorders: A Practical Guide</i> (2nd edn), Oxford: Oxford University Press, p. 291.	'Mentalising needs to be defined along the lines of being a mental process by which we attribute intentions to each other; it is how we understand each other and ourselves as being driven by underlying motives.'
2017	Bo, S., Sharp, C., Fonagy, P., and Kongerslev, M. (2017). 'Hypermentalizing, Attachment, and Epistemic Trust in Adolescent BPD: Clinical Illustrations.' <i>Personality Disorders</i> , 8(2): 172–182, p. 173.	'Mentalizing is related to the concepts of Theory of Mind, social cognition, metacognition, and emotional sensitivity, but considered a broader concept referring to the process of understanding and linking behavior and mental states'
2017	Bevington, D., Fuggle, P., Cracknell, L. and Fonagy, P. (2017). <i>Adaptive Mentalisation-Based Integrative Treatment: A Guide for Teams to Develop Systems of Care</i> , Oxford: Oxford University Press, p. 24.	'Mentalising will be considered as a mental process that is fundamentally about the constant, concerned effort required to hold an adaptive balance between competing components of psychological functioning.'
2017	Perroud, N., Badoud, D., Weibel, S., Nicastrò, R., Hasler, R., Küng, A.L., Luyten, P., Fonagy, P., Dayer, A., Aubry, J.-M., Prada, P. and Debbané, M. (2017). 'Mentalization in Adults with Attention Deficit Hyperactivity Disorder: Comparison with Controls and Patients with Borderline Personality Disorder.' <i>Psychiatry Research</i> , 256: 334–341, p. 334.	'The capacity to appreciate the mental states that underlie behaviour'
2019	Duschinsky, R., Collver, J., and Carel, H. (2019). 'Trust Comes From a Sense of Feeling One's Self Understood by Another Mind: An Interview With Peter Fonagy.' <i>Psychoanalytic Psychology</i> , 36(3): 224–227.	'To have mentalizing you need to be balanced on four dimensions of mentalizing: between cognition and affect, self and other, inside and outside, and reflective and intuitive. We concluded that if you are balanced on these four dimensions then you're mentalizing.'
2019	Bateman, A. and Fonagy, P. (2019). 'Introduction.' In Anthony Bateman and Peter Fonagy (eds), <i>Handbook of Mentalising in Mental Health Practice</i> (pp. 3–20). New York: American Psychiatric Association, p. 3.	'Mentalising describes a particular facet of the human imagination: an individual's awareness of mental states in himself or herself and in other people, particularly in explaining their actions.'

from the past. Contrary to the operationalization of reflective function (see Chapter 3), there is no explicit indication that future experiences are considered relevant for mentalizing. The reason for this is unclear but may be because concrete particulars are less available for future than present or past perceptual experience, resulting in a tangible object for mentalizing.

iii. Not simply a focus: At times, mentalization is defined as a ‘focus’ on mental states, for instance in 2008. Here mentalization is not a capacity, but the extent of sustained attention. Certainly, possession of the capacity increases the extent it will be used. But focus of mental states seems a potential correlate of mentalization, rather than a necessary element. Relevant here, perhaps, is that measures of mind-mindedness are unrelated to measures of mentalization.⁶⁴ It is telling that the definition of mentalization in terms of a focus on mental states drops away as the concept of ‘hypermentalization’—overactive and ineffective mentalization—is introduced and distinguished from ‘mentalization’ proper.

iv. Self and others: The 2012 definitions appear to distinguish the kind of mentalizing required for understanding and interpreting the mental states of oneself and visible others. The mental states of oneself require objectification and interpretation (‘seeing ourselves from the outside’); the mental states of others require inference and interpretation (‘seeing others from the inside’). However, placed together with the other definitions, such as the 2009 definition by Slade, it would appear that mentalization also still encompasses the ability to think about the self’s own social behaviour as well as the internal experience of others. These are all called ‘mentalizing,’ even if they may have different underpinning processes.

v. Intentions: The content of ‘mental states’ is variously defined. In 1995, mentalization is equated with the capacity to take multiple points of view, suggesting that thinking about mental states is the same as the ability to take multiple perspectives. However, in 1996, the content of mental states is specified as motivations/intentions. In 2000, this content is specified as ‘thoughts and feelings,’ which are presumed to be the basis for motivation. This is aligned with Target’s definition where intentions are distinguished from mental states, but recognition of both is required. In 2007, mental states are described as including ‘needs, desires, feelings, beliefs, goals, and reasons,’ which encompasses various forms of thought, feeling, and motivation. The 2009 definition gives intentions as a species of mental state, alongside feelings and thoughts, and some others. By contrast, the 2016 definition reverts to the 1996 sense of mental states as motivations/intentions, highlighting the instability of the theorized content of mental states.

Yet what is meant by intentions? Fonagy and colleagues seem to use the term in two senses. First, ‘the intentionality of mental states such as beliefs and desires refers to their “aboutness” (Brentano 1874; Dennett and Haugeland 1987)—thus a belief is “about” an actual or possible state of affairs.’⁶⁵ Following Brentano, then, for mental states to be intentional means only that they are directed towards some object, that ‘in perception something is perceived, in imagination, something is imagined, in a statement something stated, in love something loved, in

⁶⁴ E.g. Barreto, A. L., Fearon, R. P., Osório, A., Meins, E., and Martins, C. (2016). ‘Are Adult Mentalizing Abilities Associated with Mind-Mindedness?’ *International Journal of Behavioral Development*, 40(4): 296–301.

⁶⁵ Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). *Affect Regulation, Mentalisation and the Development of the Self*, London: Karnac Books, p. 147, discussing Brentano, F. ([1874] 2015) *Psychology from an Empirical Standpoint*, London: Routledge: ‘every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction towards an object’ (p. 92).

hate hated, in desire, desired, etc.⁶⁶ Brentano's meaning appears to have been much closer to our notion of *attention* than *intention*.⁶⁷ However, Dennett (mis)read into Brentano's appeal to 'intention' the implication that we interpret minds through ascription of motivations in order to account for and explain behaviour and experience, with motivations conceptualized as like individual will.⁶⁸ Fonagy and colleagues do not appear to recognize that 'intention' is being used in different ways by Brentano and Dennett. As a result, it remained unclear for Fonagy and colleagues what exactly is the object of mentalizing: i) intentional mental states, ii) mental states *and* intentions, or iii) mental states *or* intentions. The difference is potentially vast, depending on what is meant by intentional, and consequential both for theory and clinical work to encourage mentalizing. Not least it runs the risk of confusing the way a person takes a stance on, and responds to, the prompts of the state of affairs that comprises their environment (Brentano) with the expression of an individual's will (Dennett). We will return to this problem later (Chapter 9). For now it can be identified, across the definitions of mentalizing, that the general idea appears to have been that mentalization entails the acknowledgement of mental states, where mental states are implicated in motivations and intentions.

As the table of definitions of mentalizing signals, there are many meanings of mentalization. It is unlikely that any one definition can, or perhaps should, completely encompass them all. However, in considering its different definitions, a workable characterization of the core elements of mentalization can be picked out:

A capacity to:

1. conceive of
2. and make available for reconsideration
3. the thoughts
4. and feelings
5. implicated in motivations and intentions
6. in order to account for and explain
7. the observable social behaviour
8. and present and past perceptual experience
9. of oneself
10. and others.

When all 10 of these elements are present, *full mentalizing* can be regarded as present. Bringing together all elements in this configuration is rare. More commonly, *specific mentalizing* can be identified where certain elements are not needed in a particular situation. For a form of *specific mentalizing* to take place, examination of the definitions suggests that certain elements are essential. When mentalizing is solely automatic, conceiving (1) is more important than reconsidering, though the availability for reconsideration (2) remains. When mentalizing is controlled, reconsideration is dominant, though access to conceiving is still needed. In principle, either (3) (thoughts) or (4) (feelings), is absolutely required, because these are the denotation of 'mental states', though naturally the distinction between thoughts

⁶⁶ Husserl, E. ([1901] 1970). *Logical Investigations*, Volume 2, London: Routledge, p. 95, glossing his teacher Brentano's discussion of 'aboutness'.

⁶⁷ Cf. Watzl, S. (2011). 'Attention as Structuring of the Stream of Consciousness', in C. Mole, D. Smithies, and W. Wu (eds). *Attention: Philosophical and Psychological Essays*, New York: Oxford University Press, pp. 145–173.

⁶⁸ Dennett, D. C. (1991), *The Intentional Stance*, Cambridge, MA: The MIT Press.

and feeling is an abstraction, and other mental states are possible. For instance, values may be encompassed by the concept of mental states, and considered fed by both thoughts and feelings—though the definitions offer little guidance in considering such a case.

It is possible to conceive of motivations or intentions without attributing them to mental states, but this would not be mentalizing (see Chapter 5). Also essential are (2) and (6), the ability to reconsider mental states and to use them to account for or explain. Just to conceive of mental states is not enough, as shown by the rectification of Fonagy's original 1989 definition in 1995 by Fonagy and Target to 'the capacity to think about mental states'. To an extent, both (7) (observable behaviour) and (8) (present and past experience) are needed as the objects of the activities of accounting for and explaining. When one or the other is missing, the interpretation starts to spin loose and feed on itself, rather than remaining informed by the thoughts and feelings implicated in behaviour. It could be imagined that other objects of mentalization would be possible besides observable behaviour, present experience, and past experience. As we have seen, for example, Dunn proposed the social standards of groups as a possible object of mentalization. However, this possibility appears to have been excluded by Fonagy and Target. As such, examining the definitions, mentalization needs to be of either (9) (the self) or (10) (another person), or both.

So, for instance, only thoughts (3) or feelings (4) may be relevant to a particular occasion for mentalizing. This corresponds to Fonagy and Luyten's distinction between affective and cognitive mentalizing. Or, to take another example, only the mental states of oneself (9) or another (10) may be relevant to a particular situation. This corresponds to Fonagy and Luyten's distinction between mentalizing self and other. Mentalizing may also be used to conceive of thoughts and/or feelings without an attempt to then develop an account for motivations and intentions. When a therapist offers 'brief, accurate and simple statements of the analyst's perception of the patient's current mental state', this may—or may not—also entail an account of the patient's motivations or intentions.⁶⁹

Examination of the 10 elements of mentalization also revises and clarifies Fonagy and Luyten's distinction between internal and external mentalizing. Fonagy and Luyten characterized attention to (8) as 'internal mentalizing' and attention to (7) as 'external mentalizing'. But (7) alone is not mentalizing. The analysis above identifies either or both of (3) and (4) as required. Yet an important distinction can be drawn between attempts to mentalize the observable behaviour of others—and attempts to mentalize the observable behaviour of one-self. In fact, this would be a fruitful distinction because the 2012 Asen and Fonagy definition implies that these occur through distinct processes. But the implication is left unexplored, perhaps in part because it cross-cuts Fonagy and Luyten's internal/external polarity and the self/other polarity.

In their 2016 book, Bateman and Fonagy assert that 'good mentalizing takes but one form.'⁷⁰ They give two characteristics of all good mentalizing:

- 'Tentativeness—on the whole, a lack of absolute certainty about what is right and what is wrong, and a preference for complexity and relativism.

⁶⁹ Fonagy, P. and Target, M. (1995). 'Understanding the Violent Patient: The Use of the Body and the Role of the Father.' *The International Journal of Psychoanalysis*, 76: 487–501, 498. In Bion's terms, the analyst may offer awareness and comment on a patient's mental state without moving from receptive attention into explanatory forms of thought (i.e. 'downwards' on the grid). Bion, W. R. (1963). *Elements of Psycho-Analysis*, London: Karnac Books.

⁷⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 116.

- Moderation—a balanced attitude to most statements about mental states in both one-self and others that comes from accepting the possibility that one is not in a privileged position in regard to either one's own mental state or that of another person.⁷¹

They then give eight characteristics of good mentalizing of others, which differ from eight characteristics of good mentalization of oneself, suggesting that they do not take the same form.⁷²

Good mentalization of others is described as:

- acknowledging the opacity of other minds;
- not being paranoid;
- expressing a desire to reflect on how others think;
- awareness of different perspectives;
- interest in others;
- openness to discovery;
- a forgiving stance;
- oriented by a sense that, on the whole, the reactions of others are predictable given knowledge of what they think and feel.

Good mentalization of oneself is described as:

- acknowledging that the self can change;
- adopting a developmental perspective;
- recognizing that feelings can be confusing;
- recognizing that there may be feelings outside of conscious awareness
- awareness that we can have incompatible ideas and feelings;
- inquisitive about ourselves;
- interested in the way our mind works differently to others;
- recognizing the impact of affects.

Clearly the qualities in each set may vary independently of one another. One person's mentalizing of themselves may have strengths in acknowledging change over time but struggle to recognize the impact of affects. Another person's mentalizing of themselves may have strengths in inquisitiveness about their own mind and how it differs from others. But the individual may regard feelings as simply present or absent, rather than acknowledging the role of hybrids and half-formed emotions. The more important point, though, is that good mentalization of others and good mentalization of the self look very different

⁷¹ Ibid. 118.

⁷² Bateman and Fonagy also identified three qualities of the self-representation that facilitate mentalizing:

- a. Advanced pedagogic and listening skills.
- b. Autobiographical continuity—the capacity to remember oneself as a child and evidence the experience of a continuity of ideas.
- c. Rich internal life—the person rarely experiences their mind as empty or content-less' Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 117.

However, these three appear rather a hodgepodge. The first seems less a quality of the self-representation than a skill, and of more relevance for mentalizing others (depending on what is meant by pedagogic) than mentalizing the self. It is not clear exactly how autobiographical continuity facilitates mentalizing others, though access to episodic memory will certainly facilitate mentalizing the self. The construct of 'rich inner life' seems, from other statements by the authors, to be an effect of mentalizing, rather than a quality of it.

in Bateman and Fonagy's estimation. Good mentalization of others requires no developmental perspective or particular interest in feelings or their conflict. Instead, the focus is on awareness of different perspectives, relatively opaque to one another. By contrast, good mentalization of the self is not, perhaps surprisingly, characterized by a sense that we are predictable on the basis of knowledge of what we think and feel. Self-forgiveness is also not characteristic. Instead, the focus is on taking a developmental perspective and awareness of the preconscious mind. Luyten, Malcorps, Fonagy, and Ensink have subsequently added 'belief in changeability' as characteristic of good mentalizing of oneself, whereas this is not mentioned for mentalizing of others.⁷³

Overall, the qualities that characterize good mentalizing of others seem to privilege considering their different thoughts and perspectives; the qualities that characterize good mentalizing of oneself seem to privilege scrutiny of feelings and awareness of change in perspectives over time. It seems difficult to sustain Bateman and Fonagy's claim that good mentalizing 'takes but one form', when their own descriptions are considered closely. It is an encouraging development that the adaptation of the reflective functioning scale for children by Karin Ensink, in her doctoral work supervised by Target and Fonagy, distinguishes between mentalizing the self (RF-S) and mentalizing others (RF-O), and research using this scale has begun to document their distinct correlates and developmental trajectories.⁷⁴ It can be hoped that this measurement innovation will be extended to other measures of mentalization, including for reflective function as measured in adults.⁷⁵ Indeed, Debbané and Bateman and Morken and colleagues⁷⁶ have all argued that disentangling RF-S and RF-O as distinct causal processes will be important for supporting future advances in MBT.⁷⁷

⁷³ An updated list of the qualities of good mentalizing of oneself is provided by Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Assessment of Mentalising', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 37–62:

- 'Developmental perspective
- Realistic scepticism
- Internal conflict awareness
- Self-inquisitive stance
- Awareness of importance of affect
- Acknowledgment of unconscious and preconscious functioning
- Belief in changeability' (p. 46).

Comparison of the two lists indicates that interest in how our mind works differently from others has dropped out of good mentalizing, as has recognition that feelings can be confusing. They have instead been replaced by belief in changeability.

⁷⁴ Ensink, K. (2004). *Assessing Theory of Mind, Affective Understanding and Reflective Functioning in Primary School-Aged Children*. Unpublished doctoral dissertation, London: University College London; Ensink, K., Normandin, L., Target, M., Fonagy, P., Sabourin, S., and Berthelot, N. (2015). 'Mentalization in Children and Mothers in the Context of Trauma: An Initial Study of the Validity of the Child Reflective Functioning Scale'. *British Journal of Developmental Psychology*, 33(2), 203–217; Bizzi, F., Ensink, K., Borelli, J. L., Mora, S. C., and Cavanna, D. (2019). 'Attachment and Reflective Functioning in Children with Somatic Symptom Disorders and Disruptive Behavior Disorders'. *European Child & Adolescent Psychiatry*, 28(5): 705–717. A parallel distinction has been made in scaling the Parent Development Interview for mentalization by Suchman, N. E., DeCoste, C., Leigh, D., and Borelli, J. (2010). 'Reflective Functioning in Mothers with Drug Use Disorders: Implications for Dyadic Interactions with Infants and Toddlers'. *Attachment & Human Development*, 12: 567–585.

⁷⁵ For one attempt, see Dimitrijević, A., Hanak, N., Altaras Dimitrijević, A., and Jolić Marjanović, Z. (2018). 'The Mentalization Scale (MentS): A Self-Report Measure for the Assessment of Mentalizing Capacity'. *Journal of Personality Assessment*, 100(3): 268–280.

⁷⁶ Debbané, M. and Bateman, A. (2019). 'Psychosis', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 417–429, p. 426.

⁷⁷ Morken, K. T., Binder, P. E., Arefjord, N., and Karterud, S. (2019). 'Juggling Thoughts and Feelings: How Do Female Patients with Borderline Symptomology and Substance Use Disorder Experience Change in Mentalization-Based Treatment?' *Psychotherapy Research*, 29(2): 251–266.

5

Forms of non-mentalizing

Introduction

According to the definitions discussed in Chapter 4, mentalization is absent when an individual interprets their own or others' psychological experience and/or social behaviour without reference to mental states (thoughts and/or feelings). This could occur for any number of reasons. Fonagy and colleagues fail to draw an important distinction between non-mentalizing and 'not mentalizing'. In the former, mentalizing is specifically blocked. In the latter, mentalizing is neither active nor blocked. Ordinary life is full of 'not mentalizing'. Indeed, a great deal of our time is spent in interactions shaped by social conventions about the limits on appeal to mental states, where breach of these limits is treated as threatening, over-sincere, or an imposition.¹ Such conventions can help procedural interactions run smoothly, without the threat of exposure or unwanted intimacy. However, individuals who have been trained by circumstances or have trained themselves to hold fast to mentalizing, rather than finding it a capacity to turn on or off as needed, may find such interactions hard to sustain. In considering 'not mentalizing', it is important to identify that even apparently mentalizing interactional frames can be conventionalized into 'not mentalizing': 'How are you?' 'Good thanks, how are you?'

In some cases, the opening question may function as a conventionalized opening for a mentalizing interaction that may or may not be taken up.² In other cases, there is no true opening, only convention. Numerous cases of 'not mentalizing' may reflect domains of activity in which mental states are of secondary causal concern, but not blocked. Making music, for example, may not in the first instance represent a concern with mental states. Yet, attention to mental states is not blocked and may be specifically cultivated through taking an inquisitive stance to features such as sequence, tone, level of arousal, and the relation established with the audience.³ To take another example: an endocrinologist may interpret an adolescent's psychological experience through reference to hormones rather than mental states. Or a scientist may interpret an adolescent participant's experience through reference to neurological structures, or some other relatively material cause. This is not mentalizing rather than non-mentalizing. It could be contrasted to the same interpretation by a parent or a scientist who uses a material interpretation to duck consideration of the adolescent's thoughts or feelings. Then we have non-mentalizing, through a focus on physical *at the expense of* psychological explanation, even if the physical cause is causally important in shaping behaviour.

¹ See e.g. Benwell, B. and McCreaddie, M. (2016). 'Keeping "Small Talk" Small in Health-Care Encounters: Negotiating the Boundaries between On-and Off-Task Talk'. *Research on Language and Social Interaction*, 49(3): 258–271.

² Cameron, B. L. (2004). 'Ethical Moments in Practice: The Nursing "How Are You?" Revisited'. *Nursing Ethics*, 11(1): 53–62.

³ Strehlow, G. and Hannibal, N. (2019). 'Mentalizing in Improvisational Music Therapy'. *Nordic Journal of Music Therapy*, 28(4): 333–346.

Fonagy and colleagues have argued that all of us find ourselves in non-mentalizing states at points, which are simply part of the human condition. The researchers propose that non-mentalizing modes of processing operate according to very different assumptions compared to mentalized experience, and an individual in one of these states often disbelieves in mentalized accounts of reality. This can contribute to the intractability of these states, their contribution to extreme and unmeasured decision making, to conflicts in interaction with family and friends, and to the wider culture and political context of a society.⁴ Fonagy and colleagues have described three forms of non-mentalizing: pretend mode, psychic equivalence, and teleological mode. Sometimes they have implied that these are just three kinds of non-mentalizing that they have happened to stumble upon, and they have no particular importance beyond other forms of non-mentalizing, which could readily be identified.⁵ More frequently, however, pretend mode, psychic equivalence, and teleological mode are described definitively as ‘*the*’ forms of non-mentalizing.⁶ Pretend mode, psychic equivalence, and teleological mode have featured primarily in the writings of Fonagy and colleagues for clinical audiences. The characterization of the three modes of non-mentalizing is among the most potentially insightful contributions of their work. Nonetheless, no attempt has been made to operationalize the specific modes of non-mentalizing, with the exception of hypermentalizing, leaving their reflections speculative and obstructing further scientific development of the theory.⁷

This chapter will begin by describing pretend mode and specifying how it differs from mentalizing. A specific kind of pretend mode will then be described, termed ‘hypermentalizing’ by Fonagy and colleagues. Particular attention will be given to the pretend mode, in part because it has been discussed in most detail by Fonagy and colleagues; in part because it has two different forms; and also because it illustrates the way that the same capacity for pretend can be an asset or a disadvantage depending on its integration with other psychological processes and depending on the context. The chapter will then move on

⁴ Asen, E. and Fonagy, P. (2012). ‘Mentalization-Based Therapeutic Interventions for Families.’ *Journal of Family Therapy*, 34(4): 347–370.

⁵ E.g. Brent, B. K. and Fonagy, P. (2014). ‘A Mentalization-Based Treatment Approach to Disturbances of Social Understanding in Schizophrenia’, in P. Lysaker, G. Dimaggio, and M. Brüne (eds), *Social Cognition and Metacognition in Schizophrenia: Psychopathology and Treatment Approaches*, San Diego, CA: Elsevier, pp. 245–259. ‘One of the consequences of impaired mentalization from an MBT Standpoint is the likelihood that early, “pre-mentalistic” forms of thinking may emerge during stress; including: (1) psychic equivalence thinking, in which the ability to consider outside perspectives on one’s inner experience becomes lost and one’s own thoughts and/or feelings are taken as unequivocally real; (2) pretend mode thinking’ (p. 248, italics added).

⁶ Fonagy, P. and Target, M. (2004). ‘Playing with the Reality of Analytic Love: Commentary on Paper by Jody Messler Davies “Falling in Love with Love”’. *Psychoanalytic Dialogues*, 14(4): 503–515. ‘The modes of mental functioning that developmentally predate mentalization: the teleological stance, psychic equivalence, and pretend mode’ (p. 511); Fonagy, P., Campbell, C. and Bateman, A. (2017). ‘Mentalizing, Attachment, and Epistemic Trust in Group Therapy’. *International Journal of Group Psychotherapy*, 67(2): 176–201. ‘When mentalizing fails, individuals often fall back on non-mentalizing ways of behaving, which have some parallels to the ways young children behave before they have developed their full mentalizing capacities. The modes are: psychic equivalence, teleological, and pretend modes’ (p. 179).

⁷ The Reflective Functioning Questionnaire operationalizes ‘hypomentalising’, which is anticipated as the outcome of psychic equivalence and teleological mode. Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., ... and Lowyck, B. (2016). ‘Development and Validation of a Self-Report Measure of Mentalizing: The Reflective Functioning Questionnaire’. *PLoS One*, 11(7): e0158678. Some aspects of non-mentalizing were distinguished by Hausberg, M. C., Schulz, H., Piegler, T., Happach, C. G., Klöpper, M., Brütt, A. L., ... and Andreas, S. (2012). ‘Is a Self-Rated Instrument Appropriate to Assess Mentalization in Patients with Mental Disorders? Development and First Validation of the Mentalization Questionnaire (MZQ)’. *Psychotherapy Research*, 22(6): 699–709. However, Fonagy and colleagues have never cited this measure; the reason for this is unknown. A new clinician-report measure of non-mentalizing distinguishes the three modes: Gagliardini, G. and Colli, A. (2019). ‘Assessing Mentalization: Development and Preliminary Validation of the Modes of Mentalization Scale’. *Psychoanalytic Psychology*, 36(3): 249–258.

Table 5.1 Descriptions of non-mentalizing forms

Form of non-mentalizing	Description	Aspect of mentalizing
Pretend mode	To conceive of thoughts and feelings implicated in motivations or intentions	Conceiving of the thoughts and feelings implicated in motivations and intentions of oneself and others.
Pretend mode: hypermentalizing	To conceive of thoughts and feelings implicated in motivations or intentions in order to account for and explain the observable social behaviour of others	Conceiving of the thoughts and feelings implicated in motivations and intentions in order to account for and explain the observable social behaviour and others.
Psychic equivalence	To account for and explain thoughts and feelings and observable social behaviour in terms of immediate experience	Accounting for and explaining the thoughts and feelings and the observable social behaviour [in terms of] present and past perceptual experience of oneself.
Teleological mode	Observable social behaviour is used to account for and explain the motivations and intentions of others and of oneself	Observable social behaviour [are used] to account for and explain the motivations and intentions of oneself and others.

to address the two other kinds of mentalizing: psychic equivalence and teleological mode. It will be demonstrated how the forms of non-mentalizing are not simply the absence of mentalizing, but draw on specific aspects of the capacity to mentalize. This can be seen in Table 5.1, which offers a summary characterization of the key conclusions from this chapter about the forms of non-mentalizing and their respective relationships with mentalization.

Pretend mode

As described in Chapter 2, one of Fonagy's first projects after appointment as research director at the Anna Freud Centre was collaboration with Target, Edgcombe, and Miller in developing a manual for child analysis. In the course of this work, Fonagy and Target worked out 'a heuristic map of the emergence of mentalisation'. They observed that children:

appear to be able to use the notion of mental states but paradoxically use it only when they can clearly separate it from physical reality (for example, in play). In this state of mind, which we have called pretend mode, thoughts and feelings can be envisioned and talked about but they correspond to nothing real.⁸

The concept of pretend mode—together with the idea of psychic equivalence—was introduced and elaborated in a series of major papers by Fonagy and Target published in the mid-1990s in the *International Journal of Psychoanalysis* under the title 'Playing with Reality'. The

⁸ Fonagy, P. and Bateman, A. (2009). 'A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice', in Brownescombe Heller, M., and Pollet, S. (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176, p. 160–161.

concept was then further articulated by Bateman and Fonagy in their book on *Mentalisation-Based Treatment for Personality Disorders*.⁹

In the paper first introducing the concept of pretend mode, Fonagy gives the illustration of his work with a child patient. Rebecca was five years old when she was brought to the Anna Freud Centre for treatment by her mother. She lived sometimes with her grandparents and sometimes with her mother; her father had been a casual partner of her mother, and had no involvement in raising Rebecca. Rebecca's symptoms included a combination of fears and conduct problems, which began following the death of her grandfather. Her play with Fonagy consisted essentially of a single game, in which Rebecca was an older girl, Hannah, and Fonagy was cast in the role of the father. Fonagy characterized this game as pretend mode on the basis that Rebecca could not tolerate any discussion that implied questioning whether the narrative corresponded to reality. The whole question was forbidden: 'She refused to accept that her wish for grandfather or me to be her real father did not make it true. I could be either her analyst or her father, but never the analyst representing the father.'¹⁰ What made the game Rebecca was playing 'pretend mode' was that elements of imagination were being used to create accounts of her own mind and the mind of father figures specifically at the expense of her personal experience of the 'hurt and shame of being fatherless'.¹¹

In part, this use of pretend mode reflected Rebecca's developmental stage, in which the distinction between representation and reality is not always fully secured. However, Rebecca's intolerance of any questioning of her narratives of having a father also had roots in her home environment, in which her family could not discuss Rebecca's worries about lacking a father. As a consequence, Fonagy proposed, the worries were cut off from the specifics of her own life as they entered into a fantasy narrative. And they were also cut off from modulation or doubt. Furthermore, Rebecca seemed stuck in this state, stating and re-stating the narrative about having a father such that it inhibited her capacity to engage in other forms of play. Had Rebecca been able to enter into pretend mode and then exit in a flexible way, then pretend mode would not have been at the expense of tolerating thoughts and feelings. In turn, these thoughts and feelings would then have been available for modulation, rather than expressed as fears and aggression.

Fonagy and colleagues described how components of mentalizing are 'hijacked into the pretend mode of experience'.¹² This claim becomes somewhat clearer by considering the ideas of Fonagy and colleagues as responding to those of Bion. For Bion, the identification, modulation, and coordinated expression of thoughts and feelings depend on a psychological apparatus that he terms the 'alpha function'. The alpha function protects thoughts and feelings from invasion by proto-thoughts and feelings (e.g. unconscious wishes, passing ideas, aspects of embodied sensory experience) that would otherwise hinder their coherence. It also protects proto-thoughts and feelings from conscious processes that might otherwise stifle their creativity and richness, and in turn the capacity to convey this creativity and richness to determinate thoughts and feelings as the ability to consider different perspectives and develop these perspectives by learning from experience. However, the alpha function can

⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press.

¹⁰ Fonagy, P. (1995). 'Playing with Reality: The Development of Psychic Reality and its Malfunction in Borderline Personalities'. *The International Journal of Psychoanalysis*, 76: 39–44, p. 41.

¹¹ Ibid. 43.

¹² Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 437.

be fragmented back into its constituent parts. When this happens, like a vehicle, the alpha function goes into reverse.¹³ The most potent obstacle to coherent thinking and feeling is the activity of these fragments. Not only does the unintegrated activity of these fragments systematically hinder their use as part of the alpha function for coherent thinking and feeling, but there is also the potential for invasion of thoughts and feelings by proto-thoughts and feelings, and vice versa. This leads to fragmented conscious thoughts and feelings, and over-concrete proto-thoughts and feelings. Both states hinder the potential for both perspective taking and for learning from new experiences.

Likewise, for Fonagy and colleagues, the forms of non-mentalizing are so potent precisely because they ‘hijack’ the constituent parts of mentalizing. Non-mentalizing is not then simply the obverse of mentalizing, but *turns the equipment of mentalizing against itself*. This point aligns with Bion’s earlier proposal that it may be productive for theorists to shuttle between attention to the elements that facilitate learning from experience, and attention to the elements that block learning from experience.¹⁴ In the previous chapter, we considered the elements of mentalization through attention to the various definitions that have been offered by Fonagy and colleagues. Mentalization was conceptualized as a capacity to conceive of and make available for reconsideration thoughts and feelings implicated in motivations and intentions in order to account for and explain the observable social behaviour and perceptual experience of oneself and others. In these terms, pretend mode can be regarded as *conceiving of thoughts and feelings implicated in motivations or intentions*. But it is the severed first part of mentalization. For mentalization requires also some attempt to account for and explain observable social behaviour and/or perceptual experience. Also highly conspicuous in its absence is the modulated and generative doubt about thoughts and feelings that underpins reconsideration. The resulting experience has something of an ‘as if’ quality, but it is not interrogated in terms of whether it agrees with reality.

Hypermentalizing

For Fonagy, the idea of ‘as if’ thinking seemed to very much characterize his experience as a clinician working with patients with borderline personality disorder (BPD) symptoms, where all kinds of things would be narrated, but without connection to the particulars of actual experience or behaviour, and without attention to the extent to which they corresponded with reality.¹⁵ As a result, therapeutic progress grinds to a halt.

Luyten, Fonagy, Lemma, and Target suggested five features of pretend mode that distinguish it from mentalizing:

¹³ Bion, W. R. (1963). *Elements of Psycho-Analysis*, London: Karnac Books.

¹⁴ Bion, W. R. (1962). *Learning from Experience*, London: Karnac Books: ‘Instead of sense impressions being changed into alpha-elements for use in dream thoughts and unconscious waking thinking, the development of the contact-barrier is replaced by its destruction. This is effected by the reversal of alpha-function so that the contact-barrier and the dream thoughts and unconscious waking thinking which are the texture of the contact-barrier are turned into alpha-elements, divested of all characteristics that separate them from beta-elements and are then projected thus forming the beta-screen ... The reversal of alpha-function does violence to the structure associated with alpha-function’ (p. 25).

¹⁵ Fonagy, P. (1997). ‘Where Cure was Inconceivable. The Aims of Modern Psycho-Analysis with Borderline Patients’. *Texte*, 3(17): 11–25: ‘The patient’s pseudo-insights do not correspond to the primary level of their internal experience’ (p. 22). See also Fonagy, P. (2018). ‘Fonagy on “Playing with Reality: I. Theory of Mind and the Normal Development of Psychic Reality”’. *PEP/UCL Top Authors Project*, 1(1): 27.

1. First, accounts organized by pretend mode tend to be ‘overly analytical, repetitive and lengthy’.¹⁶ Without the need to lock in pragmatically to the work of attempting to account for and explain the particulars of observable social behaviour, present perceptions or concrete past experiences, the activity of conceiving of thoughts and feelings can run without restraint.
2. Second, narratives tend to be out of touch with the speaker’s own affective experience as formed by the obstacles and tasks that prompt this experience. There is an ‘as if’ quality to the discourse: cogs are spinning, but the mechanism has not engaged with practical problems and their particularity.
3. In parallel, the accounts tend to be out of touch with the genuine bases on which the speaker could claim self-worth and perceived control over their life.
4. Fourth, speech in pretend mode rather than mentalizing is signalled by difficulties for the speaker in switching perspectives—for instance, from a focus on the motivations of the self to those of the other, or vice versa. Finally, Luyten, Fonagy, Lemma, and Target proposed that accounts in pretend mode tend to position the audience in such a way as to be subject to the speaker—for example, obligated to offer attention or compassion, or to place the audience under the speaker’s control. Pretend mode is not informed by concern with the present or past perceptual experience of others, and therefore tends to treat the needs of the speaker as of utmost importance.
5. This is not to imply that mentalization is always motivated by compassionate impulses.¹⁷ Nonetheless, Luyten, Fonagy, Lemma, and Target do seem to imply that effective mentalizing should, in general, contribute to motivations more accurately reflecting the needs of others and oneself, which in turn should contribute to beneficence. (We will return to this assumption in Chapter 8).

The form of pretend mode most easily confused with mentalization has been termed by Sharp, Fonagy, and colleagues ‘hypermentalization’.¹⁸ Bateman and Fonagy define hypermentalizing as ‘a tendency to elaborate models of internal states in the absence of relevant evidence’.¹⁹ In hypermentalization, pretend mode is applied to the sustained attention to the thoughts and feelings of other people through the interpretation of their observed

¹⁶ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). ‘Depression’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, 385–418, p. 410.

¹⁷ Fonagy and colleagues offer few discussions of what motivations especially prompt mentalizing. This can give the impression that mentalizing others is free from uncivilized motivations and fantasies, such as the wish to control others, envy of their thoughts or feelings, or inability to be alone with our own thoughts and feelings. There is no reason that their position should imply this, and it would appear an implausible stance. Fonagy’s citations of Hegel are always of the *Phenomenology of Spirit*. However, a shift in Hegel’s position from the *Phenomenology* to later work such as *The Science of Logic* was towards the claim that reflection is not self-grounded, and that forms of accurate and well-judged understanding and reasoning about ourselves and others need to be examined for the forces and desires that organize and animate them, and which may not be fully understood or well reasoned. Hyppolite, J. ([1952] 1997). *Logic and Existence*, New York: SUNY University Press. Similar points have been made about reflection within psychoanalysis. See e.g. Bion, W. ([1977] 2019). *Bion in New York and Sao Paulo, and Three Tavistock Seminars*, London: Karnac Books.

¹⁸ Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., and Fonagy, P. (2011). ‘Theory of Mind and Emotion Regulation Difficulties in Adolescents with Borderline Traits’. *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(6): 563–573; Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: ‘The biggest challenge in recognising mentalising is being able to distinguish it from pseudomentalising’ (p. 127).

¹⁹ Bateman, A. and Fonagy, P. (2019). ‘Preface’, in *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, p. xx.

behaviour. The researchers characterized three forms of hypermentalization. These illustrate both the diversity of hypermentalization, and its characteristic continuities:

Intrusive pseudomentalising: opaqueness of minds is not respected; extends knowledge of thoughts and feelings beyond a specific context; presents knowledge of thoughts and feelings in an unqualified way; presents thoughts and feelings with a richness and complexity that is unlikely to be based on evidence; when challenged, defaults to nonmentalising accounts.

Overactive form of pseudomentalising: idealisation of insight for its own sake; thoughts about other felt by them as confusing and obscure.

Destructively inaccurate pseudomentalising: denial of objective realities that undermine subjective experience; cast in terms of accusations; denying someone's real feelings and replacing them with a false construction.²⁰

If pretend mode entails conceiving of thoughts and feelings implicated in motivations or intentions, hypermentalization has further components of actual mentalizing. To hypermentalize is to *conceive of thoughts and feelings implicated in motivations or intentions in order to account for and explain the observable social behaviour of others*. So close to mentalization in many regards. All that is missing are the ability for reconsideration of mental states and attention to the specifics of present or past perceptual experience. But this is enough to unmoor the other component parts of mentalization, which start to float downriver. Conclusions are generated for which there is no basis in experience, and these are insulated from meta-cognitive reappraisal.²¹ In this sense, the distinction between psychological theory and hypermentalizing is somewhat porous, and will depend on the extent to which the theory is grounded in the concrete specifics of observation and measurement, and the extent to which the theory is open to reconsideration of its constituent elements. These are qualities that Fonagy described in the 1990s and 2000s as particularly under threat in psychoanalytic theory, but which are constitutive risks of any theory-building, which must stretch out from the heavy particularities of concrete experience, making use of both the buoyancy and the gravity of concepts. Localized access to pretend mode is vivifying for innovation, nonconformity to existing common sense and ethical vision, all components to varying degrees of any adequate psychological theory. However, the generation of theory needs to be pared with its pruning and evaluation (see Chapter 2).

Bo, Sharp, Fonagy, and Kongerslev have described hypermentalization as a breakdown of communication between affective and cognitive mentalizing. We would offer a different characterization: that hypermentalization lacks components 2 (reconsideration) and 8 (present or past perceptual experience) from the definition of mentalization presented in the previous chapter. The advantage of this conceptualization is that it aligns with Bateman and Fonagy's observation that, in hypermentalization, it is controlled mentalization specifically that is offline, rather than cognition *in general*.²² Cognition in the form of component

²⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 128.

²¹ Bo, S., Sharp, C., Fonagy, P., and Kongerslev, M. (2017). 'Hypermentalizing, Attachment, and Epistemic Trust in Adolescent BPD: Clinical illustrations'. *Personality Disorders*, 8(2): 172–182.

²² Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 12–13.

6 (accounting for and explain) is still very firmly in play in hypermentalization. The other advantage of our characterization is that it is not clear what it means for a breakdown of communication to occur between affective and cognitive mentalizing. Our characterization specifies that the issue lies in the fact that reflection is not informed by perceptual experience and is not available for reconsideration, which is why it is a form of pretend mode thinking. This also explains why hypermentalization is a form of pretend mode thinking, a position which is asserted but not explained by Fonagy and colleagues.

Fonagy has speculated that the patients who helped Freud generate psychoanalysis as a talking cure were displaying symptoms of BPD, and that much of their thinking and speech reflected hypermentalizing. They could elaborate on their mental states at length. But they were also highly and quite enduringly suggestible, since their specific difficulty was in recognizing and reflecting on present or past experience as a source of thoughts and feelings. In this state, a shared pretend mode can readily be elaborated on the basis of clinician suggestion.²³ For Fonagy, a significant proportion of psychoanalytic theorizing, based on work with hypermentalizing patients, is itself the speculative attribution of mental states cut rather free of attention to concrete particulars. There are also weak mechanisms available for disconfirming psychoanalytic hypotheses, both in the consulting room and in interaction between analyst peers (see Chapter 2). The characterization of some—though by no means all—psychoanalytic theorizing as hypermentalizing appears rather a devastatingly apt characterization, from Freud to the present day.²⁴ Fonagy and colleagues have also proposed that the apparent success of patients with BPD on measures of theory of mind and interpretation of facial cues is actually because these measures fail to distinguish mentalization from hypermentalization. Individuals engaged in hypermentalizing may pay detailed attention to the external indicators of mental states assessed by these measures, but the initial ideas of participants are left unchecked by the potential for reconsideration.

To empirically examine this proposition, Sharp, Fonagy, and colleagues conducted a film-based assessment of social cognition with 111 adolescent inpatients between the ages of 12 and 17.²⁵ The movie comprised a 15-minute film about four characters at a dinner party. At 45 points the movie was stopped, and participants were asked what the characters

²³ Ezrati, O. (2014). 'Freud Off: Giving New Meaning to Psychoanalysis,' Haaretz, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>: 'From the descriptions, Freud's patients sound very borderline. Of course, it's not surprising that much of psychoanalysis was discovered with those patients, because they are incredibly accessible to elaborating their mental states. They hyper-mentalize quite a lot. Also, unfortunately, they are very suggestible. If you tell them that this or that has happened, because they can't turn inside and check out what's going on in their own mind.'

²⁴ One illustration is Freud's appeal to symbolism to 'overcome' the opacity of the patient's mind. Freud, S. ([1900] 2001). *The Interpretation of Dreams. The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 4, London: Vintage: 'We are thus obliged, in dealing with those elements of the dream-content which must be recognized as symbolic, to adopt a combined technique, which on the one hand rests on the dreamer's associations and on the other hand fills the gaps from the interpreter's knowledge of symbols' (p. 353). In itself, reference to cultural discourses in interpreting mental states seems readily justifiable (see Chapter 9). However, the very speculative account of mythic symbolism deployed by Freud can be seen to predispose hypermentalizing in his clinical work. See e.g. Kris, A. O. (1994). 'Freud's Treatment of a Narcissistic Patient.' *The International Journal of Psychoanalysis*, 75: 649–664; Koellreuter, A. (ed.), (2016). *What is this Professor Freud Like?: A Diary of an Analysis with Historical Comments*, London: Karnac Books.

²⁵ Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., and Fonagy, P. (2011). 'Theory of Mind and Emotion Regulation Difficulties in Adolescents With Borderline Traits.' *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(6): 563–573. See also Somma, A., Ferrara, M., Terrinoni, A., Frau, C., Ardizzone, I., Sharp, C., and Fossati, A. (2019). 'Hypermentalizing as a Marker of Borderline Personality Disorder in Italian Adolescents: A Cross-Cultural Replication of Sharp and Colleagues' (2011) Findings.' *Borderline Personality Disorder and Emotion Dysregulation*, 6(1): 5.

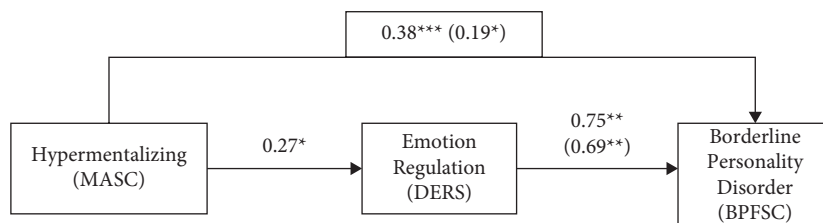


Figure 5.1 Difficulties in emotion regulation (DERS) as a mediator of the relation between hypermentalizing (MASC) and borderline personality traits (BPFSC).

Note: Values on each path are standardized β values (path coefficient). Coefficient inside parentheses are standardized partial regression coefficient from equations that include both variables with direct effects on the criterion or dependent variable. * $p < .05$; ** $p < .01$; *** $p < .001$.

Source: Reproduced from Carla Sharp et al., 'Theory of Mind and Emotion Regulation Difficulties in Adolescents With Borderline Traits', *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(6): 563–573, Figure 1, DOI: <https://doi.org/10.1016/j.jaac.2011.01.017> Copyright © 2011 American Academy of Child and Adolescent Psychiatry. Published by Elsevier Inc. All rights reserved.

were thinking, feeling, or intending. The assessment was coded to distinguish between mentalizing and hypermentalizing. A clinical interview was also used to assess for BPD, and 23% of the sample met diagnostic criteria. Participants also completed a self-report measure of emotion regulation, the 'Difficulties in Emotion Regulation Strategies Scale'. This measure treats as a single dimension a variety of somewhat different phenomena under the umbrella of 'emotion regulation': 'awareness and understanding of emotions, acceptance of emotions, the ability to engage in goal-directed behavior and refrain from impulsive behavior when experiencing negative emotions, as well as the flexible use of situationally appropriate strategies to modulate emotional responses'.²⁶

Sharp and colleagues found that the degree of mentalizing had no association with symptoms of BPD—in direct conflict with Fonagy's early theory of the developmental pathway to BPD. However, in clear agreement with Fonagy and colleagues' later account, the researchers found a very substantial ($r = .41$) relationship between hypermentalizing and symptoms of BPD, independent of age, gender, and other mental health symptoms. Sharp and colleagues also reported that 44% of the relationship between hypermentalizing and symptoms of BPD was mediated by difficulties with emotion regulation (see Figure 5.1). The findings demonstrate the importance of distinguishing between hypermentalizing and mentalizing, both for assessments of social cognition and theoretically. And the findings suggest the importance of emotion dysregulation for the stabilization of hypermentalization as a form of non-mentalizing into the symptoms associated with BPD. However, what aspects of emotion dysregulation are serving this role as yet remains unclear, due to the use of an umbrella measure of emotion regulation in this exploratory research.

²⁶ Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., and Fonagy, P. (2011). 'Theory of Mind and Emotion Regulation Difficulties in Adolescents With Borderline Traits'. *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(6): 563–573, p. 566. Later research has found that the different phenomena included in the Difficulties in Emotion Regulation Strategies Scale have a bifactor structure, with one general emotion dysregulation factor and five uncorrelated specific factors. Both the general factor and the specific factors each demonstrated incremental validity. Hallion, L. S., Steinman, S. A., Tolin, D. F., and Diefenbach, G. J. (2018). 'Psychometric Properties of the Difficulties in Emotion Regulation Scale (DERS) and its Short Forms in Adults with Emotional Disorders'. *Frontiers in Psychology*, 9: 539.

The capacity for pretend as an asset

The basis for pretend mode, hypermentalizing, and for mentalizing lies in the imaginative capacity to conceive of mental states. This capacity has been situated as resting developmentally on ‘playful exploration’, characterized by Bateman and Fonagy as a ‘basic emotion’, like anger, disgust, fear, sadness, and surprise.²⁷ They argued that the capacity for playful exploration is inherent to humans. Though Bateman and Fonagy focus their attention on symbolic play, in principle, playful exploration is much wider than this. It entails the manipulation of something in the world—for instance, a toy—to consider its modalities—for instance, what it must, should, can, may, or might do. With the maturation of a child’s capacity for cognition, playful exploration is elaborated into the ability to take pleasure and interest from play with the modalities of representations, what they must, should, can, may, or might do. Fonagy and colleagues proposed that the ability to decouple the embodied experience of observed reality, and its felt immediacy, from representations of this reality is fed by playful exploration. This ability is of critical importance for mentalizing and reflective function: it is the capacity for imaginative, counter-factual representation that allows thoughts and feelings to be conceived of in others.²⁸ Fonagy and colleagues also argued, perhaps more speculatively, that the identification and representation of thoughts and feelings in oneself likewise depend on the same process.²⁹

The imaginative capacity to decouple experience and representation allows an individual to negotiate the implications of past or present perception, somewhat or wholly de-realizing them. In his essay, ‘Psychoses and child care’, Winnicott described one form of imaginary play that he regarded as quite problematic. It is characterized by ‘the lack of a beginning and end to the game, by the degree of magical control, by the lack of organization of play material according to any one pattern, and by the inexhaustibility of the child’.³⁰ This seems to resonate with Fonagy’s account of Hannah’s narratives about her missing father, discussed above. By contrast, Winnicott describes a healthy form of de-realization, in which reality is incorporated in such a way as to make it available for perspective-taking. In agreement with Winnicott, Fonagy and colleagues proposed that the exercise of imagination can contribute to mentalization in the symbolic play of young children.

This was illustrated by a study by Tessier, Normandin, Ensink, and Fonagy.³¹ The researchers studied 39 children aged between 3 and 8 who had experienced sexual abuse, and

²⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 305. Though not cited, given the use of the term ‘exploration’, in the background here may be Bowlby and Hinde’s idea of the exploratory behavioural system as, like attachment, a form of behaviour predisposed in humans by our evolutionary history. See Bowlby, J. (1969/1982). *Attachment*, London: Penguin, p. 238 and Hinde, R. A. (1954). ‘Factors Governing the Changes in Strength Of a Partially Inborn Response. I. The Nature of the Response, and an Examination of its Course’. *Proceedings of the Royal Society, B*, 142: 306–331.

²⁸ Fonagy, P. (1991). ‘Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient’. *The International Journal of Psychoanalysis*, 72: 639–656: ‘Pretence and the understanding of another’s mental state have in common the need to be able to entertain a belief whilst at the same time knowing this to be false’; Luyten, P., Campbell, C., and Fonagy, P. (2020). ‘Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach’. *Journal of Personality*, 88(1): 88–105: ‘Mentalizing is the basic social cognitive tool that individuals use to constantly update and adaptively sketch out their imagined selves and the imagined minds of others’.

²⁹ This proposal is elaborated in a recent doctoral project, supervised by Liz Allison: Hardy, A. (2018). *The ‘Reality Oriented’ Imagination: A Philosophical Examination of the Imagination in ‘Mentalization’ and ‘Neuropsychological Analysis’*. Unpublished doctoral thesis, London: University College London.

³⁰ Winnicott, D. W. (1953). ‘Psychoses and Child Care’. *British Journal of Medical Psychology*, 26(1): 68–74.

³¹ Tessier, V. P., Normandin, L., Ensink, K., and Fonagy, P. (2016). ‘Fact or Fiction? A Longitudinal Study of Play and the Development of Reflective Functioning’. *Bulletin of the Menninger Clinic*, 80(1): 60–79.

21 matched controls who had not experienced sexual abuse. The children were assessed with the Children's Play Therapy Instrument to examine the coherence of beginning-middle-end sequences during a child's fantasy play. The children also completed the adapted Adult Attachment Interview, coded using Ensink's revision of the reflective functioning scale to be applicable to children and to distinguish mentalizing the self and mentalizing others (see Chapter 4). The researchers found that sexual abuse had a substantial association with reduced capacity to mentalize other people in the interview. But this relationship was wholly mediated by the coherence of sequences in the child's fantasy play. There was, however, no mediation by the coherence of symbolic play on the effects of sexual abuse on the abused children's reduced scores on the scale for mentalization of the self. The researchers argued that the capacity for effective symbolic play could serve to buffer the consequences of sexual abuse for making sense of other minds. But they were sorry to see that, at least on the evidence of this study, such buffering did not extend to the child's capacity for identifying and interpreting their own mental states in interview.

The capacity for imagination is important not just for children but for adults too. Fonagy and Target described in detail the value of play and imagination for sensitive caregiving informed by reflective functioning. As we saw in Chapter 3, for Ainsworth, sensitivity requires more than the caregiver accurately identifying the child's physical behaviour—for instance, an infant's cry. Sensitivity requires the caregiver to 'perceive and to interpret accurately the signals and communications implicit in her infant's behaviour'.³² This entails conceiving of invisible mental states, in this case feelings implicated in motivations, which already in infancy may be as varied as hunger, discomfort, pain, loneliness, boredom, or confusion. To Ainsworth's account of sensitivity, Fonagy and Target added that reflective function entails not just a capacity to acknowledge the child's mental states, but also the capacity for perspective-taking, a kind of play with representations. Perspective-taking allows the caregiver to put their child's demands into context and not be overwhelmed by them: a difficult task when confronted with an infant's seemingly endless crying. Nonetheless, Fonagy and Target proposed that if the caregiver feels overwhelmed and totally subjected to the child's displayed emotion, this in turn offers reduced containment for the child.³³ By contrast, if a child's emotion is met with acknowledgement but also with the caregiver's capacity to hold this in perspective, to remain capable of imagination and perspective-taking, then this will contribute to the child's ability to 'ultimately use the parent's representation of his internal reality as the seed for his own symbolic thought, his representation of his own representations'.³⁴

Fonagy and colleagues have claimed that imagination and spontaneity can support innovation, nonconformity, and ethical vision. In fact, engaging in psychological theory itself entails use of this imaginative capacity. Problems with imagination and spontaneity arise, they have argued, when the 'gating mechanism' that would usually 'bridle' these responses fails. Imagination and spontaneity are then cut loose from i) the concrete particulars of

³² Ainsworth, M. (1969). 'Sensitivity vs. Insensitivity to the Baby's Signals Scale.' Accessed at: http://www.psychology.sunysb.edu/attachment/measures/content/ainsworth_scales.html.

³³ Target, M. and Fonagy, P. (1996). 'Playing with Reality II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479. See also Byrne, G., Slead, M., Midgley, N., Fearon, R. M. P., Mein, C., Bateman, A., and Fonagy, P. (2019). 'Lighthouse Parenting Programme: Description and Pilot Evaluation of Mentalization-Based Treatment (MBT) to Address Child Maltreatment'. *Clinical Child Psychology and Psychiatry*, 24(4): 680–693.

³⁴ Target, M. and Fonagy, P. (1996). 'Playing with Reality II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479, p. 472.

experienced reality and/or ii) from the stabilizing effects of social referencing and norms.³⁵ Their emphasis on this point can sometimes give the impression that Fonagy and colleagues value the stabilizing effects of social referencing and framing over nonconformity. This is probably an effect of the fact that their work emerges in part from a clinical context, in which stability and common reference points are important proximal assets and goals. However, Fonagy and colleagues have at times demonstrated a wider distrust in the judgement of individuals separated from social referencing and norms: 'in general when people are left on their own to figure things out they tend to go astray in their thoughts and feelings.'³⁶

Nonetheless, in their writings on play Fonagy and Target strongly underscored the benefits of a child's developmental capacity for decoupling representation from past or present experience. They reviewed evidence that suggested that the subtlety of a child's understanding of their own and others' mental states as the basis for one perspective, rather than a simple expression of reality, is *greater* while in pretend play.³⁷ The addition of playful exploration may make examination of mental states easier, in part perhaps because the heavy particularities of concrete experience can be dropped away. This could be why the use of projective doll play tasks in the assessment of children's experience of relationships or thinking can be much more effective than straightforward versions of the same tasks.³⁸

Fonagy and colleagues have argued that both the symbolic play of children and the creativity of adults are nourished by access to i) the concrete particulars of lived experience; and ii) to common reference points with others, including forms of social recognition and cultural knowledge.³⁹ Yet Fonagy and colleagues also appear to accept that creativity is vivified by the capacity—in turn—to 'gate' knowledge of concrete particulars and common reference points in the perspectives of others, and access pretend mode for particular purposes. As we saw in Chapter 3, Fonagy praised the work of creators such as Baudelaire, Rimbaud, and Nerval, for whom the bridle did not sit comfortably.⁴⁰ Given the reflections of Baudelaire,

³⁵ Fonagy, P., Allison, E., and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 63–77, p. 70; Bateman, A. and Fonagy, P. (2019). 'Preface', in *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, p. xvii. For a comparison of mentalizing with the psychoanalytic notion of the superego, see Jurist, E. L. (2014). 'Whatever Happened to the Superego?: Loewald and the Future of Psychoanalysis'. *Psychoanalytic Psychology*, 31(4): 489–501.

³⁶ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 325.

³⁷ E.g. Dias, M. G. and Harris, P. L. (1990). 'The Influence of the Imagination on Reasoning by Young Children'. *British Journal of Developmental Psychology*, 8(4): 305–318; Gopnik, A. and Slaughter, V. (1991). 'Young Children's Understanding of Changes in their Mental States'. *Child Development*, 62(1): 98–110.

³⁸ Woolgar, M., Steele, H., Steele, M., Yabsley, S., and Fonagy, P. (2001). 'Children's Play Narrative Responses to Hypothetical Dilemmas and their Awareness of Moral Emotions'. *British Journal of Developmental Psychology*, 19(1): 115–128; Steele, M., Steele, H., Woolgar, M., Yabsley, S., Fonagy, P., Johnson, D. (2003). 'An Attachment Perspective on Children's Emotion Narratives: Links cross Generations', in R. N. Emde, D. P. Wolf, and D. Oppenheim (eds), *Revealing the Inner Worlds of Young Children: The MacArthur Story Stem Battery and Parent–Child Narratives*, Oxford: Oxford University Press, pp. 163–181.

³⁹ Fonagy, P. and Allison, E. (2018). 'The Origin of Human Life: A Psychoanalytic Developmental Perspective', European Psychoanalytical Federation, 31st Annual Conference, Warsaw, 24 March: 'The potential space that Winnicott refers to is the imagined common ground between self-reflection (the experienced self) and the experience of the other's reflection on the self. The larger this common ground, the closer the shape of self and the other's perceived view of the self, the higher the level of epistemic trust experienced by the patient and the more open the patient's mind is to imagining the analyst's experience in relation to physical or mental phenomena to which the patient so far has no access. Patients unlock their imagination and accept a new reality when they feel that their reality has been adequately imagined and accepted.'

⁴⁰ It is curious that Fonagy almost never mentions authors of fiction. The same is true for Bion. Nonetheless, both cite Baudelaire and Rimbaud as authors who illustrate the role of imagination and proto-thoughts and feelings in the effective and powerful depiction of mental states. Bion, W. R. (1976). 'Emotional Turbulence', in *Clinical*

Rimbaud, and Nerval themselves on the nature and costs of creativity, it may be suspected that for these authors this 'access' was not always easy to modulate, and that this was intricately bound up with its power and effectiveness.

For instance, in Baudelaire's poem 'Crowds', he argues that both access to the commons and the potential for independence from it are necessary for genuine creation: 'multitude, solitude: equal and interchangeable terms for the active, fertile poet. The man who cannot people his solitude will not be able, either, to be alone in a busy crowd.'⁴¹ In *The Painter of Modern Life* he defied his reader to identify a truly beautiful work of cultural production that was not an effective composite of intersubjectivity and current cultural knowledge, together with something autonomous of them, out of keeping with anything timely.⁴² However, for Baudelaire, this composite was not in the first instance a methodological product of conscious intention, but of slow work on the self to become able to sustain and benefit from both rampages of individual imagination and immersion in the crowd. To take another example, Nerval's final novel, left unfinished by his suicide, cuts off after the protagonist's remarks of gratitude for the events that had 'rendered me back to the renewed affection of my family and friends', which had helped him to 'a healthier view of the world of illusions in which I had for some time lived'. Simultaneously, he offered gratitude for his journeys beyond common social and cultural reference points, for all that these journeys could sometimes resemble hell.⁴³

Besides the production of creative works, another context in which pretend mode has specific uses, one that especially interested Fonagy and Target, is the therapeutic setting. Here the relationship between pretend mode and shared perspective taking is squared in the collaboration between patient and therapist. In therapy there is a 'dual reality'.⁴⁴ This can take clients a bit of time to recognize, and early sessions are often spent helping the patient discover this. The patient must talk seriously about feelings and ideas which they also know are false, as if this was the most important thing in the world. And then leave the therapeutic setting and get on with responding to the practical challenges of a life that only partly intersect with the matters under discussion in their therapy. The transference relationship with the therapist is a kind of 'illusion'. As Fonagy puts it wryly, 'most of us who have had analytic experiences know what is meant by transference love, but few of us have serenaded our analyst on moon-lit summer nights.'⁴⁵ However, this aspect of pretend within the therapeutic relationship is not a deficit, but part of how it works. Therapy loses its effectiveness if it is wholly subjected to the ugly, dismal, unsafe, or stultifying, because 'the intrusion of unassimilable reality destroys the possibility of play, and this includes playing

Seminars and Other Works, London: Karnac Books, pp. 295–305, p. 304. This seems less likely to be influence, and more likely theoretical convergence to the point of aesthetic overlaps.

⁴¹ Baudelaire, C.-P. (1995). *Selected Poems*, trans. Carol Clark, London: Penguin, p. 201.

⁴² Baudelaire, C.-P. (1972). 'The Painter of Modern Life', in *Selected Writings on Art and Literature*, trans. P. E. Charvet, London: Penguin, pp. 390–436.

⁴³ Nerval, G. ([1855] 1999). 'Aurélia', in Richard Sieburth (ed.), *Selected Writings*, London: Penguin, pp. 265–316, p. 316.

⁴⁴ Fonagy, P. (1991). 'Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient'. *The International Journal of Psychoanalysis*, 72: 639–656, p. 652.

⁴⁵ See also Fonagy, P. (1995). '2: Peter Fonagy'. *British Journal of Psychotherapy*, 11(4): 575–584: 'Therapy, certainly psychoanalytic therapy, is in many respects a "pretend" experience. Therapist and patient discuss fantasies, feelings and ideas which they "know" at the same time to be false. John Klauber wrote beautifully about this in his paper on transference as an illusion (in Klauber and others, 1988). Most of us who have had analytic experiences know what is meant by transference love, but few of us have serenaded our analyst on moon-lit summer nights' (p. 578), citing Klauber, John (1988). *Illusion and Spontaneity in Psychoanalysis*, London: Free Association Books.

with ideas.⁴⁶ Part of the difficulty in supporting mentalization in self-critical, depressed patients is that 'these patients may have little capacity for adaptive pretend functioning', cutting off possibilities for playing with ideas.⁴⁷ Yet, when there is also some pretend in play to offer safety, therapy is most beneficial when it intersects with the patient's lived experience outside therapy and the ugly, dismal, unsafe, or stultifying things that it may contain. This was part of what impressed Fonagy about Hurry's approach to his own analysis as an adolescent: his cheap, ugly car—seen by Hurry out of the window—could be valued for what it signified to an isolated young man (see Chapter 1).

In a child's—or an adult's—symbolic narratives, observations of social behaviour and aspects of perceptual experience are incorporated into the play, giving it vitality and resonance. This is the paradox and magic of play. But in pretend mode as a form of non-mentalizing, thoughts and feelings implicated in motivations or intentions are conceived specifically at the expense of awareness of observable social behaviour and/or perceptual experience. Fonagy and Target proposed that factors that undermine the development of mentalization—in the 1990s and 2000s they emphasized especially disorganized attachment relationships—may be anticipated to hinder the integration of imagination and understanding.⁴⁸ On the one hand, the capacity for decoupling representation may risk representation uninformed by reality. This will result in chaotic or confusing narratives without effective sequencing. On the other hand, when mentalizing capacities are fragile, the decoupled representation may still *feel* like reality. Then children's experiences of pretend may be quite anxiety-provoking.⁴⁹ Recently, Fonagy and colleagues have come to characterize dissociation as an extreme form of pretend mode, in which observable social reality and/or perceptual experience are not simply ignored, but lost from awareness, through immersion in the pretend mode of non-mentalizing.⁵⁰ Pretend mode does not simply predispose mental health problems, but in fact is proposed as the process through which certain symptoms such as dissociation occur.

Fonagy and Target have offered guidance on helping children and adults integrate imagination and understanding in the context of therapeutic work. Greater use of pretend mode is developmentally appropriate for children. Here the therapist may wish to facilitate pretend, while nourishing it with concrete particulars and intersubjective understanding, allowing the child to conceive of and reconsider their experiences. Target has reflected that, when working with older patients who lean heavily on pretend mode, in her experience, clinical 'technique needed to be modified in the early stages to enable these patients, already too loosely connected to reality, to engage more authentically. Consistent gentle confrontation, and little development of a shared pretend mode, helped to create traction. After that, a more usual analytic technique became effective.'⁵¹ Skårderud and Fonagy have also stressed

⁴⁶ Target, M. (1998). 'The Recovered Memories Controversy'. *International Journal of Psychoanalysis*, 79: 1015–1028, p. 1026.

⁴⁷ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 385–418, p. 411.

⁴⁸ Target, M. and Fonagy, P. (1996). 'Playing with Reality II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479.

⁴⁹ Kaplan, N. R. (1989). *Individual Differences in Six-Year-Olds' Thoughts about Separation: Predicted from Attachment to Mother at One Year*. Unpublished doctoral thesis, Berkeley: University of California. Luyten, P. and Fonagy, P. (2014). 'Assessing Mentalizing in Attachment Contexts', in S. Farnfield and P. Holmes (eds), *The Routledge Handbook of Attachment: Assessment*, London: Routledge, pp. 210–216.

⁵⁰ Target, M. (2007). 'The Interface between Attachment and Intersubjectivity: Another Contribution from Karlen Lyons-Ruth'. *Psychoanalytic Inquiry*, 26(4): 617–621, p. 620.

⁵¹ Target, M. (2008). 'Some Thoughts on Lying and Pretending'. Plenary delivered to the British Psychoanalytical Society Scientific Meeting, 18 June 2008, unpublished manuscript, Mary Target personal archive.

the importance of authenticity for work with adolescent and adult patients reliant on pretend mode processing. They urge that clinicians feel able to think aloud, referring to ideas, perspectives, meanings, feelings. But that these should always be tagged as belonging to the therapist. Otherwise there is a risk that the patient will pick up these elements as the basis for a shared pretend mode, without traction for the concretely lived experience of the patient and the personal and social challenges they face.⁵²

Psychic equivalence

Besides pretend mode, a second kind of non-mentalizing described by Fonagy and colleagues was what they termed 'psychic equivalence'. In this state, aspects of present experience are taken to represent facts that are inevitable and true.⁵³ Within psychic equivalence, it is extremely difficult for the individual to entertain the idea that there could be any other perspective on how things are besides how it feels right now. As we saw in Chapter 4, mentalization may be conceptualized as a capacity to conceive of and make available for reconsideration thoughts and feelings implicated in motivations and intentions in order to account for and explain the observable social behaviour and perceptual experience of oneself and others. In these terms, *psychic equivalence means to account for and explain thoughts and feelings and observable social behaviour in terms of an immediate experience*. Missing here is the capacity to conceive of and reconsider perceptions of things. Present perceptions, even sensory perceptions, are profoundly mediated by wider psychological and social processes. But they may possess the quality of *feeling* immediate to the individual. Hegel termed this the feeling of 'sense certainty'.⁵⁴ In psychic equivalence, sense certainty is mistaken for reality. Present perceptions and their felt immediacy dominate the scene. For instance, feeling betrayed, an individual interprets the mental states and behaviour of themselves and others in terms of a betrayal. Feeling a sense of togetherness with another person, psychic equivalence leads an individual to the conclusion that this is because they *are* in reality together or a single unit, which can set up future disappointments.

Rather than perceptual experience being understood in terms of feelings and thoughts as implicated in motivations and intentions, within psychic equivalence mental states and behaviour are understood only in terms of an immediate experience. Indeed, even the capacity

⁵² Skårderud, F. and Fonagy, P. (2012). 'Eating disorders', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Publishing, pp. 347–384, p. 378.

⁵³ Initially there was some ambiguity in Fonagy and Target's account of psychic equivalence about what inner experience was considered 'equivalent' to. They later acknowledged that 'In previous writings, we had conflated two features of psychic equivalence: (a) equation with other minds and (b) equation with the physical'. Fonagy, P. and Target, M. (2007). 'Playing with Reality: IV. A Theory of External Reality Rooted in Intersubjectivity'. *International Journal of Psychoanalysis*, 88: 917–937, p. 927. This ambiguity is resolved by the 2010s, where—following the main line of argument present already in Fonagy and Hepworth's writings from the 1990s—psychic equivalence was taken to be an equation of personal perception and reality. What matters for psychic equivalence is the lack of mediation that would occur when perceptions are recognized as perceptions, rather than reality.

⁵⁴ Hegel, G. W. F. ([1807] 1977). *The Phenomenology of Spirit*, trans. A.V. Miller, Oxford: Oxford University Press. Heidegger offers a characterization of sense certainty that highlights the way that sense certainty undercuts perspective-taking since there is no perspective available to take: 'sense certainty asserts itself to be the unchanging relation, admitting no distinction between I and object, wherein the poles of the relation as well as the relation itself are undifferentiated and not distinguished.' This requires ongoing efforts to dodge or suppress information that might otherwise suggest this distinction: 'What is shown again and again is the effort to keep from falling out of immediacy ... and not posing questions.' Heidegger, M. ([1930–1931] 1988). *Hegel's Phenomenology of Spirit*, trans. Parvis Emad and Kenneth Maly, Indianapolis: Indiana University Press, pp. 70–2.

to compare and integrate information from different senses may be lost, as any immediate experience, even from one sense alone, gains the quality of sense certainty and comes to feel absolute and necessary. Possible differences between past experience and current perceptions are likewise lost or sidelined.

In clinical settings, psychic equivalence can contribute to a variety of symptoms. One of the central axes of trauma symptomatology is flashbacks, in which memories are experienced in the mode of psychic equivalence, feeling terrifyingly immediate, actual, and inevitable.⁵⁵ Another of trauma's central axes is avoidance of reminders of the event. Again, this can be interpreted as the operation of psychic equivalence, in which the experience of the reminder is felt to threaten the immediacy of the traumatic experience. The feeling of concreteness associated with psychic equivalence can also contribute to somatic symptoms following trauma, through the equation of psychological and physical pain and exhaustion.⁵⁶ In contrast to experiences that feel over-real and inescapable, prompted by psychic equivalence, 'the alternative to this mode of functioning, the heir of the pretend mode of psychic reality, is dissociation in the wake of trauma. The most characteristic feature of traumatization is the oscillation between these two modes of experiencing internal reality.'⁵⁷

Psychic equivalence can also contribute to the symptoms of BPD in the domain of personal relationships. Under the terms of psychic equivalence, individuals cannot experience their fears, grievances, expectations, and other affectively laden attributions as unfounded, even when they seem improbable. And the present appears as the only reality: the sense of time, of existing with a past and future, is disturbed by the immediacy and unalterable quality of the individual's experience. This hinders access to strategies such as patience, helpful distraction, taking a moment to regain composure, and other techniques deployed by the 'gating mechanism' that would otherwise prevent our current perceptions from stampeding into the form of set certainties.⁵⁸ In a state of psychic equivalence, individuals can also be hyper-reactive to shame, because the badness of the self is experienced as total and without possible respite.⁵⁹ Likewise, frustrations can blow up into dramatic outbursts and aggression, but without an accurate sense of the cause of the frustration: 'In explaining a violent outburst, for instance, the patient might refer to the oppressive character of the room he was in.'⁶⁰

Fonagy and colleagues proposed that the most important trigger for psychic equivalence is high arousal/emotionality. Strong emotion can also help sustain psychic equivalence by hindering access to controlled mentalizing.⁶¹ Fonagy and colleagues do not, so far as we

⁵⁵ Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 442.

⁵⁶ Luyten, P., Fonagy, P., Lemma, A. and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 385–418, p. 397.

⁵⁷ Fonagy, P., Target, M., Gergely, G., Allen, J. G. and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 442.

⁵⁸ Fonagy, P., Allison, E. and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 70.

⁵⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 70.

⁶⁰ Ibid. 130.

⁶¹ Ibid. 43.

can tell, ascribe a fundamental proximal trigger for pretend mode or teleological mode in this way, putting into relief the special link they see between high arousal/emotionality and psychic equivalence. In fact, given that they regard high arousal/emotionality as giving experience intensity, immediacy, and urgency, there is a sense in which the boundary between the constructs of high arousal/emotionality and psychic equivalence is permeable. Bateman, Fonagy, and Campbell propose that ‘emotions always feel real.’⁶² Nonetheless, not all high arousal/emotionality leads to psychic equivalence. When Fonagy and colleagues discuss the kind of high arousal/emotionality that leads to psychic equivalence, they tend to use the term ‘dysregulated emotion,’⁶³ which might give one axis of specification to this notoriously hazy concept.⁶⁴

As we saw earlier, Fonagy and colleagues do not wish to denigrate non-mentalizing, despite the fact that they feel that it can contribute to suffering, hinder measured decision making, and fuel interpersonal conflict. The capacity for pretend and ‘as if’ thinking can be an asset. Likewise, Fonagy and colleagues observe that the raw certainty and immediacy of psychic equivalence plays a part in many of the most dramatic, invigorating, and meaningful aspects of human life. Psychic equivalence can provide essential winter fuel for the hard times faced by individuals holding fast to ethical conviction or sense of vocation, allowing a personal perception of injustice or calling to outweigh other concerns. This can allow such values to take precedence over mental states such as disappointment or a wish for easier, pragmatic solutions. Likewise, psychic equivalence can contribute to the feeling of our personal relationships as consequential, and even to some of their satisfactions. Part of the pleasure of adult sexuality, Fonagy and Target claim, stems from entry into a state of psychic equivalence: the immediate experience of giving and receiving pleasure can

⁶² Bateman, A., Fonagy, P. and Campbell, C. (2019). ‘Borderline Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 323–334, p. 330. Given reflection, this claim seems somewhat unlikely. Jurist gives a whole chapter in his book on mentalizing emotion to discussing emotions that do not feel wholly real, and that prompt doubt in themselves. In a sense, this claim might instead be regarded as definitional, in clarifying what Fonagy and colleagues actually mean by emotion—i.e. an affective state which cannot prompt doubt in the reality of what is perceived. This gives further evidence for the porous boundary between the constructs of arousal/emotionality and psychic equivalence, discussed earlier in the chapter, if states of arousal/emotionality imply a lack of doubt in current perceptions. On emotions that do not feel wholly real, see Jurist, E. (2018). *Minding Emotions: Cultivating Mentalisation in Psychotherapy*, New York: Guilford Press, Chapter 1.

⁶³ E.g. Bateman, A., Fonagy, P., and Campbell, C. (2019). ‘Borderline Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 323–334: ‘We see dysregulated emotional thinking as triggering the concreteness of the psychic equivalence mode, which in turn makes it difficult for the individual to accept alternative perspectives that could help to contextualise and downregulate the intensity of the experience’ (p. 326).

⁶⁴ Dysregulation connotes both intensity and out-of-context affect, but in fact only one or the other may be a problem. It suggests all affects that are poorly regulated, but it is not clear that joy or awe or political outrage, even if dysregulated, would be what Fonagy and colleagues wish to capture. It is also not clear whether the term ‘emotion dysregulation’ is intended to capture states such as flat affect. This would seem to depend on what is meant by ‘regulation’. Berlant, L. (2015). ‘Structures of Unfeeling: Mysterious Skin’. *International Journal of Politics, Culture, and Society*, 28(3): 191–213. One well-known definition of emotion regulation is that of Gross, in which it encompasses five processes: situation selection, situation modification, attentional deployment, cognitive change, and response modulation. It might be that for Fonagy and colleagues emotion dysregulation is the breakdown of one or more of these processes and/or other social processes. However, if so, then this would call for work to clarify the specific breakdowns on these five dimensions they anticipate would be associated with non-mentalizing modes. Appeal merely to the umbrella term ‘dysregulated emotion’ has negative consequences for the work of Fonagy and colleagues—for instance, in contributing to the difficulties keeping track of attentional deployment, which is sometimes distinguished from and sometimes included within emotion dysregulation; also in contributing to difficulties distinguishing emotion dysregulation from non-mentalizing. Gross, J. J. and Jazaieri, H. (2014). ‘Emotion, Emotion Regulation, and Psychopathology: An Affective Science Perspective’. *Clinical Psychological Science*, 2(4): 387–401.

be felt to characterize the relationship and the lovers themselves (see Chapter 6).⁶⁵ The psychic equivalence underpinning romantic love, especially early in the relationship, is so taken for granted that expressions of this mode of thought are culturally mandatory and conventional. It is hard to find a Valentine's Day card that is not framed from the position of psychic equivalence: characterizing the beloved as the best person in the world; collapsing past, present, and future in asserting eternal feelings; the relationship is described as magical, miraculous, without mundane cause or precedent.

Fonagy and Target also claimed that making use of psychic equivalence is part of how talking therapies offer relief from symptoms. Talking therapies provide 'a unique opportunity for psychic change by offering a situation where an individual may re-experience the concreteness of the psychic equivalent mode of functioning with regard to his emotional life, together with the imaginative freedom and clear separation from external reality.'⁶⁶ The patient can experience fantasies, thoughts, and feelings as real and concrete, without these threatening to become so outside the clinic, thanks to the aspects of therapy that helpfully invoke pretend mode.⁶⁷ For instance, the therapist can offer reassurances that the patient's expression of a wish to do something illegal or immoral will not be taken as, in itself, cause for concern or result in negative consequences for the patient or their relationship with the therapist. Once fantasies, thoughts, and feelings in their intensity and immediacy have emerged, the therapist can then work with the patient to make these adequately specified, represented, and made available for discussion and thought. This can only take place, however, once genuinely held fantasies, thoughts, and feelings are made available: 'without the emotional immediacy and conviction, starting with the rawness of psychic equivalence, reflection would be sterile intellectualisation.'⁶⁸ It is precisely out of an experience of immediately perceived reality as the truth that reality, perception, and truth can be articulated and distinguished.

Teleological mode

In research on theory of mind in the early 1990s, cognitive scientists drew a distinction between 'teleological causality' and 'psychological causality'. In the former, an individual's action is represented as causally related to a future goal state; in the latter, action is represented as 'intentionally' related to a future goal state.

Q: Why did the chicken cross the road?

Teleological causality: To get to the other side.

Psychological causality: It wanted to be on the other side.⁶⁹

⁶⁵ Fonagy, P. and Target, M. (2004). 'Playing with the Reality of Analytic Love: Commentary on Paper by Jody Messler Davies "Falling in Love with Love"'. *Psychoanalytic Dialogues*, 14(4): 503–515; Target, M. (2007). 'Is our Sexuality our Own? A Developmental Model of Sexuality Based on Early Affect Mirroring'. *British Journal of Psychotherapy* 23(4): 517–530.

⁶⁶ Target, M. and Fonagy, P. (1996). 'Playing with Reality II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479, p. 469.

⁶⁷ Target, M. (2008). 'Some Thoughts on Lying and Pretending'. Plenary delivered to the British Psycho-Analytical Society Scientific Meeting, 18 June. Accessed at: <https://wcp.memberclicks.net/assets/documents/friday-eveplenary-target-lyingandpretending.pdf>. 'Islands of unbearable experience (as with Dr P), may be identified through the analyst's sensitivity to the rawness of psychic equivalence, and his maintenance of the safe "pretend" frame.'

⁶⁸ Ibid.

⁶⁹ Csibra, G. and Gergely, G. (1998). 'The Teleological Origins of Mentalistic Action Explanations: A Developmental Hypothesis'. *Developmental Science*, 1(2): 255–259, p. 255.

It was anticipated that teleological and psychological interpretations could be hierarchically organized. In one hierarchy, the cause of an action could be understood with reference to an 'intention': 'he thanked me *because* he was grateful'. Or, in another hierarchy, an intention could be understood in terms of an action: 'he thanked me *so must have been* grateful'.⁷⁰

Drawing on these ideas, Gergely, Csibra, and colleagues examined the emergence of the capacities for representing a person's action as causally related to a future goal state.⁷¹ In papers over the late 1990s, Gergely and Csibra characterized mentalizing explanations as those in which teleological causality occurs in the service of psychological causality. Conversely, teleological causality in the absence of concern with intentions represented a kind of non-mentalizing when applied to the explanation of social behaviour:

Developmentally, teleological interpretations are transformed into causal mentalistic ones by 'mentalizing' the explanatory constructs of the teleological stance: i.e. by turning representations of actual reality constraints into 'beliefs' (which mentally represent such constraints), and representations of future goal states of reality into 'desires' (which mentally represent goal-states).⁷²

Gergely and Csibra argued that the two stances have very different implications for social interaction. When teleological causality alone is dominant, action can be modified only through direct and prompt action—for instance, through physical obstruction. Yet, when intentions are used to account for causal actions, 'one becomes able to influence the other's actions by changing the mental causes that generate it: i.e. by modifying the other's inferred beliefs and desires through communicative interventions such as informing, promising, or pleading'.⁷³

Fonagy and Target accepted the claim that teleological thinking in the absence of concern with intentions represented a kind of non-mentalizing.⁷⁴ To pretend mode and psychic equivalence, they therefore added a third kind of non-mentalizing, which they termed 'teleological mode'. This they defined as a form of mental processing in which behaviour is interpreted and evaluated only for its observable consequences without reconsideration of the mental states that may have motivated or prompted it. As with pretend mode and psychic equivalence, Fonagy and Target regarded teleological mode as having a precursor in the thinking characteristic of young children:

The compelling nature of physical reality is also obvious when children only impute intention from what is physically apparent. We noted that this teleological mode of thinking was present from a very early stage, but is compelling for all of us at moments when mentalizing

⁷⁰ Leslie, A. M. (1993). 'A Theory of Agency', Technical Report TR-12, Rutgers University Center for Cognitive Science.

⁷¹ Gergely, G., Nadasdy, Z., Csibra, G. and Biro, S. (1995). 'Taking the Intentional Stance at 12 Months of Age'. *Cognition*, 56: 165–193.

⁷² Gergely, G. and Csibra, G. (1997). 'Teleological Reasoning in Infancy: The Infant's Naive Theory of Rational Action: A Reply to Premack and Premack'. *Cognition*, 63(2): 227–233, pp. 231–232.

⁷³ Csibra, G. and Gergely, G. (1998). 'The Teleological Origins of Mentalistic Action Explanations: A Developmental Hypothesis'. *Developmental Science*, 1(2): 255–259, p. 259.

⁷⁴ First mentioned in Fonagy, P. (2000). 'Attachment and Borderline Personality Disorder'. *Journal of the American Psychoanalytic Association*, 48(4): 1129–1146.

has ceased, when physical reassurance is demanded and required if emotion regulation is to be reinstated.⁷⁵

Likewise, in ‘states of reverie, dreaming, or intense emotional arousal, these old “vestigial” structures may become reactivated’, and the world’s meanings become oriented around the physical qualities of things rather than thoughts and feelings about them.⁷⁶ For instance, unmentalized feelings of inner disintegration and mess evoking intense emotional arousal may prompt a bout of cleaning or list-making.⁷⁷ The gesture of organizing is forced to carry a burden of achieving felt coherence, since this cannot be shouldered by acknowledged feelings. Risky sexual behaviour may likewise have a basis in many cases in teleological mode—for instance, in the assumption that interpersonal affection or self-worth can be real only if accompanied by physical behaviour taken to demonstrate this.⁷⁸

This is not to say that thoughts and feelings are absent in teleological mode: mental states can be recognized, but only when they are felt to be observable.⁷⁹ They do not serve as a basis for developing inferences about motivations. Instead, thoughts and feelings are equated with concrete behaviour and its observable causes. Luyten, Fonagy, Lemma, and Target provided as an example the way that an individual may only feel loved if he or she has access to tangible, ideally physical signifiers of others’ affection.⁸⁰ Such thinking can make individuals operating in teleological mode unusually attentive and vigilant to others’ behavioural cues. This permits an indirect attention to and awareness of mental states, via ‘hypervigilance’ to their observable behaviour.⁸¹ However, mental states are not used to account for the motives behind behaviour. Nor are they available for reconsideration. As a consequence, one’s own and others’

⁷⁵ Fonagy, P. and Bateman, A. (2009). ‘A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice’, in Brownescombe Heller, M., and Pollet, S. (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176, p. 161.

⁷⁶ Fonagy, P. (1994). ‘Mental Representations from an Intergenerational Cognitive Science Perspective’. *Infant Mental Health Journal*, 15(1): 57–68, p. 64.

⁷⁷ This example is from Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 137. For careful disambiguation of the concept of enactment, which has links to teleological mode, see Bohleber, W., Fonagy, P., Jiménez, J. P., Scarfone, D., Varvin, S., and Zysman, S. (2013). ‘Towards a Better Use of Psychoanalytic Concepts: A Model Illustrated Using the Concept of Enactment’. *The International Journal of Psychoanalysis*, 94(3): 501–530.

⁷⁸ Bateman, A., Fonagy, P. and Campbell, C. (2019). ‘Borderline Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 323–334, p. 327.

⁷⁹ Fonagy, P. (2016). ‘The Role of Attachment, Epistemic Trust and Resilience in Personality Disorder: A Trans-Theoretical Reformulation’, *DMM News*, 22, September. Accessed at: <http://www.iasa-dmm.org/images/uploads/DMM%20%2322%20Sept%2016%20English.pdf>: ‘The teleological mode refers to states of mind where mental attitudes are only recognised if they are accompanied by a tangible signifier and lead to a definite outcome. Hence, the individual can recognize the existence and potential importance of states of mind, but this recognition is limited to very concrete, observable situations’ (p. 3).

⁸⁰ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). ‘Depression’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 385–418, p. 400.

⁸¹ Fonagy, P. (2000). ‘Attachment and Borderline Personality Disorder’. *Journal of the American Psychoanalytic Association*, 48(4): 1129–1146: ‘It is conceivable that, at a stage when such nonconscious mind-reading skills begin to evolve, the implications for the child of trying to infer the intentions behind their caregivers’ reactions are so negative that they are forced to fall back on the strategy of influencing the other by action rather than by words. However, they retain access at a nonconscious level to mental states, although they repudiate consciousness of it. It is not that borderline patients are “mind blind”; it is rather that they are not “mind conscious”. They pick up on cues that influence the behavioural system, but these do not surface in terms of conscious inferences’ (p. 1141). On the focus on observable behaviour under teleological mode as a form of hypervigilance, see Bateman, A., Fonagy, P., and Campbell, C. (2019). ‘Borderline Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 323–334, p. 328.

physical cues and observable behaviour can give diverse and conflicting signals without the means to sift and interpret them. As we saw in Chapter 1, Fónagy and Fonagy had highlighted that secondary meanings are present across human speech in both childhood and adulthood, making almost every communication resonate with a variety of interpretable meanings. Only a fraction of this is concerted and conscious: 'the face and the prosodic qualities of speech are a theatre for too many states of mind for a coherent understanding of a relationship to be reliably achieved by these means.'⁸² These external signs taken in isolation are unstable and labile, often showing contradictions between primary and secondary meanings. As a result, external behaviours can fail to reflect the underpinning, psychological organization.⁸³

Like pretend mode and psychic equivalence, teleological mode has been regarded by Fonagy and colleagues as a hijacked component of mentalization. Specifically, it represents the severed component of accounting for and explaining observable social behaviour.⁸⁴ In mentalizing a capacity to conceive of and make available for reconsideration thoughts and feelings implicated in motivations and intentions is deployed in order to achieve this accounting and explaining. In teleological mode, *observable social behaviour is used to account for and explain the motivations and intentions of others and oneself*. This is in contrast to mentalization where thoughts and feelings are used to account for and explain observable behaviour and perceptual experience. In the terms of Luyten and Fonagy's four polarities, teleological mode is a kind of non-mentalizing in which 'external mentalizing' is used to infer internal states.

An interesting finding by Gagliardini and Colli has been that mentalization-based therapists and psychoanalytic therapists may be more disposed to spot teleological mode than cognitive-behavioural therapists, given the greater focus of CBT on behaviours and manifestations of mental states, which may make teleological mode normalized to a certain extent by this modality.⁸⁵ The precise contours of teleological mode can be seen by examining a strategy often used in CBT for patients with personality disorders: drawing up contracts with patients. A contract represents a physical embodiment of a determinate plan for future observable behaviour. Use of contracts became a popular clinical strategy from the 1980s in mandating patients with BPD to put aside specific self-destructive coping strategies.⁸⁶ Luyten and Fonagy, writing in 2012, argued that use of contracts is inappropriate when patients are operating to a significant degree in teleological mode. In this mode, patients do not have adequate 'capacity for volition', so are not able to offer meaningful consent for a contract, which requires a minimal degree of knowledge of one's own plans.⁸⁷ By contrast, Skårderud and Fonagy, writing in the same year, argued that 'because of the teleological stance that those with eating disorders often display, we consider written agreements to be very useful.

⁸² Fonagy, P. (2011). 'Discussion of Juan Pablo Jimenez's Paper, "A Fundamental Dilemma of Psychoanalytic Technique. Reflections on the analysis of a Perverse Paranoid Patient"', in J. P. Jimenez and Muguillansky, R. (eds), *Clinical and Theoretical Aspects of Perversion*, London: Karnac Books, pp. 63–76, p. 71.

⁸³ The classic work here is Sroufe, L. A. and Waters, E. (1977). 'Attachment as an Organizational Construct'. *Child Development*, 48(4): 1184–1199. See also Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). *Affect Regulation, Mentalization and the Development of the Self*, London: Karnac Books.

⁸⁴ In a presentation from October 2007, Fonagy characterized teleological mode as 'misuse of (cognitive) mentalization'. Fonagy, P. (2007). 'Infancy and Borderline Personality Disorder'. Accessed at: <https://www.hebpsy.net/articles.asp?id=1510>.

⁸⁵ Gagliardini, G. and Colli, A. (2019). 'Assessing Mentalization: Development and Preliminary Validation of the Modes of Mentalization Scale'. *Psychoanalytic Psychology*, 36(3): 249–258.

⁸⁶ See Bloom, H. and Rosenbluth, M. (1989). 'The Use of Contracts in the Inpatient Treatment of the Borderline Personality Disorder'. *Psychiatric Quarterly*, 60(4): 317–327; Miller, L. J. (1990). 'The Formal Treatment Contract in the Inpatient Management of Borderline Personality Disorder'. *Psychiatric Services*, 41(9): 985–987.

⁸⁷ Luyten, P. and Fonagy, P. (2012). 'The Multidimensional Construct of Mentalization and its Relevance to Understanding Borderline Personality Disorder', in A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience*, Oxford: Oxford University Press, pp. 405–427, p. 407.

They can assist in bringing about the mental attitude that “I believe it because I see it and I try to cooperate because I actually have signed”⁸⁸

The apparent conflict is resolved, however, by Bateman and Fonagy in their 2016 book on *Mentalization-Based Treatment for Personality Disorders*. There the authors advocated against contracts that assume an individual’s capacity to know his or her own present or future intentions, and especially those that lead to discharge when the conditions are not fulfilled. However, contracts that target particular areas likely to cause problems in treatment specifically when the patient enters teleological mode can be a useful resource for engaging the patient in that mode.⁸⁹ For example, a contract that states ‘Even if I demand it, please be careful about changing my medication in a crisis. I can consider this later when I am calmer’ can be helpful in response to the teleological demand to *do something* in response to felt crisis, even if this something is not a well-measured or ultimately beneficial response.⁹⁰ The case of written contracts between therapist and patient illustrates that what is missing in teleological mode is above all the capacity to conceive of and reflect on the feelings and thoughts implicated in intentions, which can be accessed only indirectly insofar as they are equated with tangible signifiers such as a contract.

Why three forms of non-mentalizing?

Table 5.1 summarises the forms of non-mentalizing discussed so far. As mentioned at the start of this chapter, Fonagy and colleagues have tended to describe pretend mode, psychic equivalence, and teleological mode as ‘the’ three forms of non-mentalizing. Occasionally they have been referred to as three possible forms among others. And on one occasion hypermentalizing was added as a fourth.⁹¹ But in general the description of *the* three, with a definite article, is consistent. In Fonagy and Target’s initial introduction of pretend mode and psychic equivalence, the pair are presented as a natural fit as the opposite of mentalization. Mentalization was conceptualized as attention to and reflection on the reality of relationships, combined with the imaginative capacity to infer these invisible internal states. With the actual reality of relationships brought offline, Fonagy and Target assumed the result would be pretend mode. When the capacity to imagine invisible internal states went offline, Fonagy and Target assumed that the result would be psychic equivalence. In fact, before the co-authored papers with Target, Fonagy had initially termed psychic equivalence ‘actual

⁸⁸ Skårderud, F. and Fonagy, P. (2012). ‘Eating Disorders’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 347–384, p. 366. This aligns with Fonagy’s earliest discussion of contracts with patients with borderline personality disorder, in which he and Higgitt were broadly in favour. Higgitt, A. and Fonagy, P. (1992). ‘Psychotherapy in Borderline and Narcissistic Personality Disorder’. *The British Journal of Psychiatry*, 161(1): 23–43.

⁸⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: ‘Because of the teleological stance that those with eating disorders often display, we consider written agreements to be very useful here (p. 99); ‘Contracts tend to be individualised and specific, often targeting particular areas likely to cause problems in treatment. We are not great proponents of draconian contracts likely to lead to discharge when their conditions are not fulfilled. Fluctuating mentalising capacity means that a patient who agrees to a contract at one point may not actually have the same competence in a different context’ (p. 166).

⁹⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 62–64.

⁹¹ Luyten, P., Fonagy, P., Lowyck, B., and Vermote, R. (2012). ‘The Assessment of Mentalization’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 43–66, Table 2.10.

mode' to signify the over-definite sense that immediate experience is taken to represent facts that are inevitable and true. The phrase 'actual mode' signals the binary opposition he saw with pretend mode.⁹²

This model was forced to shift in the 2000s with the addition of teleological mode. Perhaps it would have shifted anyway, as in the 2000s Fonagy and Target showed growing recognition that psychic equivalence needed better specification than the absence of imagination. Psychic equivalence seemed rather to account for observed behaviour and inferred mental states in terms of current perceptions. Fonagy and Target also showed growing recognition that pretend mode needed better specification about what aspect of reality had been sequestered. Yet through the 2000s there was no attempt to offer a model of the three forms of non-mentalizing, nor to operationalize the concepts. The reason for this is difficult to identify. It would have been quite possible, for instance, to develop a coding system of the Adult Attachment Interview for identifying pretend mode, psychic equivalence, and teleological mode. In the mid-2000s, Target supervised three doctoral projects creating and evaluating new scales for the Adult Attachment Interview based on the ideas about personality disorder she, Fonagy, and colleagues had been developing. However, no attempt was made to scale the modes of non-mentalizing. And the work from the doctoral projects was ultimately left unpublished.⁹³

Ultimately, without operationalization and a clear model, the items contained in the container of 'non-mentalizing' were prone to tumble about. For instance, Fonagy and collaborators have at times reverted to their early account of a 'prementalistic psychic reality, largely split into psychic equivalence and pretend modes' as the polar opposites of one another, and in which teleological mode has no specified place.⁹⁴ Similarly, the MBT Adherence and Competence Scale, developed by Karterud and an international team including Bateman and Fonagy, specifies that mentalization-based treatment requires that the therapist identify and seek to intervene with pretend mode and psychic equivalence, but there is no requirement for an MBT therapist to address teleological mode.⁹⁵

One proposal put forward at times in the 1990s and early 2000s was that the logic of the three modes of non-mentalizing was that they reflected the thinking of pre-school children. During this time, Target was engrossed in reading the case files of child patients from the Anna Freud retrospective study, and Fonagy was practising as a child analyst. Fonagy may

⁹² Fonagy, P. (1995). 'Playing with Reality: The Development of Psychic Reality and its Malfunction in Borderline Personalities'. *The International Journal of Psychoanalysis*, 76(1): 39–44: 'A stage of development when children experience an equivalence between the internal and external. I shall call this "actual mode", alongside a representational mode of psychic functioning decoupled from external reality, a "pretend mode"' (p. 39).

⁹³ Pearson, J. (2007). *Evaluating an Alternative Coding Manual for the AAI for Use with People with Personality Disorders*. Unpublished DCLinPsy thesis, London: University College London; Lee, T. (2007). *The Development of an Alternative Personality Disorder Coding Manual for Use with the Adult Attachment Interview (AAI): A Psychoanalytical Approach*. Unpublished DCLinPsy thesis, London: University College London; Buettner, A. (2009). *Personality Pathology and Memories of Childhood: Development of a New, Psychoanalytically Informed Measure*. Unpublished doctoral thesis, London: University College London.

⁹⁴ E.g. Allen, J. G., Lemma, A., and Fonagy, P. (2012). 'Trauma', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 419–444: 'The inhibition of mentalising in a traumatising, hyperactivated attachment relationship leads to a prementalistic psychic reality, largely split into psychic equivalence and pretend modes' (p. 432).

⁹⁵ Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C., ... and Bateman, A. W. (2013). 'The MBT Adherence and Competence Scale (MBT-ACS): Development, Structure and Reliability'. *Psychotherapy Research*, 23(6): 705–717. Though it should be noted that identifying and working with psychic equivalence appears to be both rare, and difficult to code reliably, using this scale. It may be that the scale will undergo revision or refinement in the coming years. Simonsen, S., Juul, S., Kongerslev, M., Bo, S., Folmo, E. and Karterud, S. (2019). 'The Mentalization-Based Therapy Adherence and Quality Scale (MBT-AQS): Reliability in a Clinical Setting'. *Nordic Psychology*, 71(2): 104–115.

have had in mind Ivan Fónagy's observations of phenomena similar to psychic equivalence, pretend mode, and teleological mode in his diary of Peter's development (see Chapter 1). Yet, from the early 2000s, Fonagy and Target began to criticize the focus of psychoanalytic theory on early childhood developmental stages in conceptualizing the basis for adult mental illness. This entailed criticism of their own earlier work to the extent that it remained indebted to such an approach. They concluded that at best metaphorical or heuristic links could be drawn between the mental processes characteristic of children and the mental processes characteristic of unwell adults.⁹⁶ Ultimately, 'confidence in mapping particular forms of psychopathy to specific developmental epochs is unjustified ... The metaphorical "baby" of psychoanalytic theory which stands for "the past" will probably have to be abandoned.'⁹⁷ Appeal to a basis in pre-school forms of thinking could not offer a sufficient justification for pretend mode, psychic equivalence, and teleological mode as the three forms of non-mentalizing. Today, forms of non-mentalizing are still described by some of Fonagy's more psychoanalytically-oriented collaborators as "pre-mentalizing", and this is even formalized in the name of one of the scales of the Parental Reflective Functioning Questionnaire.⁹⁸ However this appears to be a holdover rather than a considered stance. The characterization of non-mentalizing as reflecting a childish form of thinking has generally been abandoned by Fonagy, even if to a lesser extent by some of his collaborators.

With Fonagy and Luyten's 2009 conceptualization of mentalization as having four poles, this introduced an account that promised not only a justification for the modes of non-mentalizing, but also an account of the specific proximal contexts in which they would occur. Since the introduction of this account, it has subsequently been advocated by Fonagy and colleagues that 'imbalances within the dimensions of mentalizing predictably generate the nonmentalizing modes': 'Psychic equivalence is inevitable if emotion (affect) dominates cognition. Teleological mode follows from an exclusive focus on external features of the neglect of the internal. Pretend mode thinking and hypermentalizing are unavoidable if reflective, explicit, controlled mentalizing is not well established'.⁹⁹ This new framework appears to offer a logic as to why only three forms of mentalizing have been identified over two decades. Mentalizing has four poles: internal and external; affective and cognitive; self and other; implicit and explicit. External mentalizing at the expense of internalizing produces teleological mode. Affective mentalizing at the expense of cognitive mentalizing produces psychic equivalence. And automatic mentalizing at the expense of explicit mentalizing produces pretend mode.¹⁰⁰

Yet this framework faces significant problems, partly inherited from trouble with Fonagy and Luyten's original scheme of oppositions. First, the absence of cognitive mentalizing

⁹⁶ Fonagy, P. (2003). 'Genetics, Developmental Psychopathology, and Psychoanalytic Theory: The Case for Ending our (Not So) Splendid Isolation'. *Psychoanalytic Inquiry*, 23(2): 218–247: "Through our growing knowledge of infants' actual capacities we may limit speculation concerning the impact of infancy on adult function. The projective processes of infancy are unlikely to work in the adultomorphic way described by Bion (1957, 1962a, b) and Klein (1935, 1952, 1958), but this does not mean that these descriptions do not contain important truths about adult mental function, simply that "infancy" is used metaphorically in these theorizations about mental process." (231).

⁹⁷ Fonagy, P. (2010). 'Development', in R. M. Skelton (ed.), *The Edinburgh International Encyclopaedia of Psychoanalysis*, Edinburgh: Edinburgh University Press, pp. 118–121, p. 120.

⁹⁸ Luyten, P., Mayes, L. C., Nijssens, L., and Fonagy, P. (2017). The parental reflective functioning questionnaire: development and preliminary validation. *PLoS ONE*, 12(5): e0176218.

⁹⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 19.

¹⁰⁰ E.g. Bateman, A. and Fonagy, P. (2019). 'Introduction', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 14.

alone does not prompt pretend mode, though it plays a major role in sustaining it. Rather, as we have seen, pretend mode also requires that mental processing is not informed by current or past experience and its concrete constraints and tasks. This was the emphasis of Fonagy and Target's early account but is missing as it has no place in Fonagy and Luyten's model. Furthermore, explicit mentalizing at the expense of internal mentalizing would still be mentalizing. It would just be ponderous. To be able to engage in automatic mentalizing is adaptive. But Fonagy and colleagues' recent definitions of mentalizing as a balance between the four poles (see Chapter 4) entail a conflation of mentalization with beneficial psychological functioning, at least in terms of the opposition between implicit and explicit mentalizing.

Second, problems also arise with the account of psychic equivalence as affective mentalizing at the expense of cognitive mentalizing. Affective mentalizing, according to Fonagy and Luyten, is the capacity to identify and interpret feelings in oneself or others. It is far from clear that psychic equivalence entails strengths in this capacity. Conversely, cognitive mentalizing was described as the capacity to identify and interpret thoughts in oneself or others. Again, as we saw earlier, psychic equivalence might better be characterized as accounting for and explaining thoughts and feelings and observable social behaviour in terms of immediate experience. Our embodied response to sensory cues means that present perceptions are the central site of affective experience.¹⁰¹ This gives the impression of affect-led non-mentalizing in psychic equivalence. But this does not therefore imply strengths in attending to or in interpreting the feelings of others, or of oneself. An individual may face systematic psychological obstacles to conceiving of or reconsidering either i) thoughts or ii) feelings in themselves or others as implicated in motivations. These predicaments should receive a name and further consideration by those interested in mentalization (not least as a basis for critical dialogue with recent developments in work on 'theory of mind'). But the key point here is that *neither* situation would necessarily imply psychic equivalence.

Such concerns suggest the need for an alternative justification for and logic to the salience of pretend mode, psychic equivalence, and teleological mode. In the previous chapter, mentalization was defined as a capacity to:

1. conceive of
2. and make available for reconsideration
3. the thoughts and
4. feelings
5. implicated in motivations and intentions
6. in order to account for and explain
7. the observable social behaviour
8. and present or past perceptual experience
9. of oneself
10. and others.

¹⁰¹ Skårderud, F. (2007). 'Eating One's Words, Part II: The Embodied Mind and Reflective Function in Anorexia Nervosa—Theory'. *European Eating Disorders Review*, 15(4): 243–252; Shai, D. and Fonagy, P. (2014). 'Beyond Words: Parental Embodied Mentalizing and the Parent Infant Dance', in M. Mikulincer and P. R. Shaver (eds), *Mechanisms of Social Connections: From Brain to Group*, Washington, DC: American Psychological Association, pp. 185–203. See also Merleau-Ponty, M. ([1945] 2012). *Phenomenology of Perception*, trans. Donald A. Landes, London: Routledge.

Exercise of this capacity can be regarded as requiring two major syntheses, which then in turn need to be integrated.¹⁰² The first synthesis is of ‘experience’. This entails *conceiving* of and *perceiving* reality, both mental and physical. Addressing this synthesis, Winnicott distinguished between the ‘subjectively conceived’ and the ‘objectively perceived’ self and other as interacting components within human perception, and urges that only when these are brought together is mental life experienced as alive and creatively usable.¹⁰³ He was especially worried by a situation he described as the ‘false self’, by which he meant a situation in which the experiences of others are used to account for and explain the motivations and intentions of others and oneself, to the exclusion of one’s own ‘subjectively conceived’ and the ‘objectively perceived’ experiences.¹⁰⁴ In interpreting the idea of ‘objectively perceived’ experience, Winnicott emphasized that an objectively perceived object is still a perception, and therefore subjective in a sense, rather than the thing-in-itself. His emphasis is how we integrate perceptions granted or imposed by the object with perceptions stemming from what we want from the object and how we find ourselves imagining it. On similar lines, mentalization depends upon the achievement of a subjective conception of other and/or oneself, fed by specific information apprehended by perceptions of felt or reported experience and observable behaviour. Drawing on Winnicott, Fonagy and colleagues have regularly emphasized that we depend upon an element of imagination in understanding motivations and intentions, which always have a certain opacity to us.¹⁰⁵

The second great synthesis is of what Fonagy and colleagues term ‘understanding’. This entails the activity of accounting for and explaining behaviour and/or experience with the

¹⁰² Drawing here from Hegel’s remarks on Kant’s synthesis of apperception and faculty of understanding. Hegel, G.W.F. ([1816] 2010). *The Science of Logic*, trans. and ed. G. Di Giovanni. Cambridge: Cambridge University Press. For Hegel, sense certainty (psychic equivalence) and mere appearance (pretend mode) were two moments in the dialectic that produces both lived experience, and the potential for consciousness of that experience (reflection).

¹⁰³ Winnicott, D. W. (1960). ‘The Theory of the Parent–Infant Relationship’. *The International Journal of Psychoanalysis*, 41: 585–595; Winnicott, D. W. (1971). *Playing and Reality*, London: Routledge.

¹⁰⁴ Dependence on the experiences and accounts of others for accounting for oneself is not in itself a false self in Winnicott’s terms. He terms this ‘compliance’, and sees it as an important developmental achievement. The false self arises when this dependence is exclusive, and one’s own experiences cannot be mined as a resource for this accounting activity. Winnicott, D. W. ([1960] 1965). *Ego Distortion in Terms of True and False Self, the Maturational Process and the Facilitating Environment*, New York: International Universities Press, pp. 140–152.

¹⁰⁵ E.g. Fonagy, P. and Target, M. (2006). ‘The Mentalization-Focused Approach To Self Pathology’. *Journal of Personality Disorders*, 20(6): 544–576: ‘Mentalizing is imaginative because we have to imagine what other people might be thinking or feeling; an important indicator of high quality of mentalization is the awareness that we cannot know absolutely what is in someone else’s mind. We suggest that a similar kind of imaginative leap is required to understand one’s own mental experience, particularly in relation to emotionally charged issues’ (p. 544). There are significant parallels here between the position of Fonagy and colleagues and that of Kant in conceptualizing ‘productive imagination’ and conceptual understanding within human cognition, and the necessity of both. Of particular relevance is §24 ‘On the application of the categories to objects of the senses in general’ in the doctrine of the elements in *Critique of Pure Reason*. In Kant’s *Lectures on Logic*, he suggested that ‘imagination and understanding are two friends who cannot do without one another but cannot stand one another either, for one always harms the other. The more universal the understanding is in its rules, the more perfect it is, but if it wants to consider things in concreto then it absolutely cannot do without the imagination.’ Kant, I. (2012). *Lectures on Logic*, trans. J. Michael Young, Cambridge: Cambridge University Press, p. 447: Fonagy and colleagues even tend, like Kant, to align productive imagination with automatic processing and conceptual understanding with controlled mentalizing, though Fonagy and colleagues are less strict (but also less clear) in this assumption of alignment than Kant. For Kant’s position, see Leech, J. (2014). ‘Making Modal Distinctions: Kant on the Possible, the Actual, and the Intuitive Understanding’. *Kantian Review*, 19(3): 339–365: ‘We need the concepts of possibility and actuality to cope with our divided cognitive architecture. There is something about the way we are able to cognize objects, come into epistemic contact with them, and think about them, that leaves a gap between our thoughts of things and the existence of those things. If we are to assure ourselves of genuine representation of real objects, and the possibility of knowledge of them, then we need a way to distinguish the actual cases from the merely possible. We at least need to be able to conceptualize this difference to make sense of sceptical worries, and to formulate strategies for avoiding error (e.g. seek out some direct experience of a purported possible thing—a consistent concept is not sufficient for existence)’ (p. 358–359).

discernment of motivations and intentions.¹⁰⁶ When conception is poorly fed by specific information, the result is pretend mode: the conception of thoughts and feelings is understood to account for them. When this is combined with the second synthesis, and used for accounting for and explaining motivations and intentions, this produces hypermentalisation. When specific information from perceived reality is not treated as a personal perception, but felt simply as reality itself, the result is a sense of overwhelming immediacy. Combined with the second synthesis, and used for accounting for and explaining motivations and intentions, the result is psychic equivalence. Here personal experience is used to account for and explain the motivations and intentions of others and oneself. Mentalization also depends upon the achievement of an account or explanation of our own or others' behaviour or experience. Teleological mode runs this process backwards, using observable social behaviour to account for and explain the motivations and intentions of others and oneself.

Consideration of mentalization in terms of the syntheses of experience and understanding helps clarify why pretend mode, psychic equivalence, and teleological mode are not just incidental obstacles to mentalizing, but fundamentally incompatible. For full or specific mentalization to be present, an individual needs to be able to understand how misunderstanding of experience can take place. Otherwise there is no access to the generative doubt that underpins inquisitive stance. This is specifically what the three modes of non-mentalizing obstruct, even if they do so each in different ways. In turn, it may be anticipated that lack of access to generative doubt helps stabilize non-mentalizing. Allison and Fonagy have recently described 'undue certainty about the veracity of an idea as psychic equivalence, while a total repudiation of this certainty we denoted as pretend mode; both are characteristic of a prementalizing phase in the development of psychic reality.'¹⁰⁷ Here, again, we can see teleological mode dropping out when Fonagy and colleagues account for the logic of non-mentalizing. However, in the same terms, teleological mode can be considered as undue certainty that the meaning of social behaviour is signified by its outcome. In all three cases, generative doubt is not possible, and so neither is the inquisitive stance about mental states and pursuit of perspective-taking. As we saw in Chapter 4, Bateman and Fonagy have situated tentativeness and moderation as qualities of all good mentalizing.¹⁰⁸

The importance of generative doubt is signalled by Fonagy's characterization from the start of mentalization as making mental states available for reconsideration, albeit with confusion about whether this required explicit representation or could be an implicit process.¹⁰⁹ The importance of generative doubt has been further headlined by recent theoretical work. Drawing on a model of resilience developed by Kalisch and colleagues,¹¹⁰ Fonagy and

¹⁰⁶ The term 'understanding' has varied meanings of course, and is not wholly stable across the writings of Fonagy and colleagues. Nonetheless, when the term appears, it tends to entail the development of an accurate account of something through appeal to intentional mental states. This has also been Fonagy's account of the concept of understanding in a recent interview: BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>. Though certainly there are differences, a predecessor for this account of 'understanding' is Jasper's definition, as a process through which 'we sink ourselves into the psychic situation and understand genetically by empathy how one psychic event emerges from another', Jaspers, K. ([1946] 1963). *General Psychopathology*, trans. J. Hoenig and M. W. Hamilton, Baltimore, MD: Johns Hopkins University Press, p. 301. Perhaps the greatest difference here is that, whereas Jaspers appeals to empathy, mentalization is conceptualized as focused on the discernment of motivations and intentions.

¹⁰⁷ Allison, E. and Fonagy, P. (2016). 'When is truth relevant?' *Psychoanalytic Quarterly*, 85(2): 275–303, p. 279.

¹⁰⁸ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 118.

¹⁰⁹ Fonagy, P. (1989). 'On Tolerating Mental States: Theory of Mind in Borderline Patients'. *Bulletin of the Anna Freud Centre*, 12: 91–115, p. 96.

¹¹⁰ Kalisch, R., Müller, M. B., and Tüscher, O. (2015). 'A Conceptual Framework for the Neurobiological Study of Resilience'. *Behavioral and Brain Sciences*, 38: e.92. See also Kalisch, R., Cramer, A. O. J., Binder, H., Fritz, J.,

collaborators have characterized two forms of reconsideration that they regard as of special importance for how mentalization contributes to resilience.¹¹¹ A first is the capacity to reconsider the valence of a present situation. Resilience is fostered by the capacity to reappraise the situation to identify those elements that may be used constructively.¹¹² Undue certainty or uncertainty are hinderances for this appraisal, which depends on a generative doubt about our first impressions.¹¹³ A second form of reconsideration that contributes to resilience is reappraisal of past experience. Past experience may press in on us with certain impressions and sensations, as well as with thoughts or feelings about what has happened. Fonagy and colleagues regard resilience as fostered by the capacity to reappraise the past to achieve accuracy in identification of present-day threats. This includes the capacity to reappraise threat-associated impressions and sensations stemming from the past, that might contribute to inaccurate and overinclusive perceptions of threat in the present. Again, generative doubt is critical to this reappraisal: undue certainty or uncertainty both hinder the potential for reconsidering the past in inquisitive stance.

A methodological development has also underlined the centrality of generative doubt and inquisitive stance in the conceptualization of mentalizing. Recently, Fonagy and colleagues have worked to develop a self-report measure of mentalizing, with the promise of reduced labour and greater scalability compared with the Adult Attachment Interview. The construction of a self-report scale is, on the surface, paradoxical. How can an individual report on their capacity for conceiving and considering mental states, if this capacity is required in order to know the answer to the question? How can pretend mode, psychic equivalence, and teleological mode be identified by an individual using them, when their effect is to make their own premises appear certain? After some trial and error, Fonagy and colleagues sought to circumvent these problems by creating a Reflective Functioning Questionnaire (RFQ)

Leertouwer, I., Lunansky, G., Meyer, B., et al. (2019). 'Deconstructing and Reconstructing Resilience: A Dynamic Network Approach'. *Perspectives on Psychological Science*, 14(5): 765–777.

¹¹¹ The term 'resilience' is a widely used buzzword. As Fonagy notes, 'resilience has been conceptualised variously as a tool, characteristic, potential, attitude, act, asset, value, process, trait, skill, resource, strength ... capacity, disposition, performance, competency', and many others. The definition preferred by Fonagy is that of Rudrauf: 'The ability of a system to resist dynamically a perturbation or adverse condition that challenges the integrity of its normal operation and to preserve function as a result in reference to some initial design or normative functional standards'. Rudrauf, D. (2014). 'Structure-Function Relationships behind the Phenomenon of Cognitive Resilience in Neurology: Insights for Neuroscience and Medicine'. *Advances in Neuroscience*: 1–28, p. 4. Fonagy, P. (2015). 'Psychotherapy for Emerging Borderline Personality Disorder'. Keynote address to the European Society of Child & Adolescent Psychiatry, Madrid, 21 June. Accessed at: <https://www.escap.eu/uploads/Events/Madrid%202015/peter-fonagy-escap-presentation.pdf>.

¹¹² Freund and Staudinger have criticized Kalisch for assuming that only positive appraisals contribute to resilience. They regard the capacity for negative appraisals as also important. Freund, A. M. and Staudinger, U. M. (2015). 'The Value of "Negative" Appraisals for Resilience. Is Positive (Re) Appraisal Always Good and Negative Always Bad?' *Behavioral and Brain Sciences*, 38: e101. If Fonagy and colleagues are read as simply endorsing Kalisch's model, they would be vulnerable to the same criticism. However, it would appear—though it is not spelt out—that they apply the model especially to reconsideration rather than the conception of experience. The criticism by Freund and Staudinger therefore does not apply. In fact, the capacity to reappraise present-day thoughts and feelings more soberly is raised in Fonagy, P., Allison, E., and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, 63–77, p. 68. Likewise implied is the contribution to resilience of a capacity to reappraise positive past thoughts and feelings, not just past threat-related sensations. These latter points are dealt with explicitly in Bo, S., Bateman, A., and Kongerslev, M. T. (2019). 'Mentalization-Based Group Therapy for Adolescents with Avoidant Personality Disorder: Adaptations and Findings from a Practice-Based Pilot Evaluation'. *Journal of Infant, Child, and Adolescent Psychotherapy*, 18(3): 249–262.

¹¹³ Badoud and colleagues have referred to experiences of self and others unhindered by undue certainty or undue uncertainty about mental states as "calibrated mentalisation". Badoud, D., Prada, P., Nicastro, R., Germond, C., Luyten, P., Perroud, N., and Debbané, M. (2018). 'Attachment and Reflective Functioning in Women with Borderline Personality Disorder'. *Journal of Personality Disorders*, 32(1): 17–30, p. 22.

with two scales.¹¹⁴ One scale represented over-certainty (RFQc) about the minds of oneself and others (e.g. 'I always know what I feel', 'I usually know exactly what other people are thinking', 'My intuition about a person is hardly ever wrong'). The other scale represented indiscriminate uncertainty (RFQu) about the minds of oneself and others (e.g. 'People's thoughts are a mystery to me', 'I don't always know why I do what I do', 'I believe there's no point trying to guess what's on someone else's mind').¹¹⁵

It should be acknowledged that an important limitation of the RFQ is that it does not seek to capture any positive qualities of mentalizing that may be distinct from the absence of non-mentalizing, such as curiosity about mental states, concern to integrate discrepant information about mental states, or effective use of support from others in conceiving or reconsidering mental states.¹¹⁶ The Parental Reflective Functioning Questionnaire, developed subsequent to the Reflective Functioning Questionnaire, does have three scales, one of which is 'interest and curiosity' in the child's mental state, which seems an attempt to capture some of the positive content of mentalizing irreducible to the absence of non-mentalizing.¹¹⁷ Nonetheless, the use of well-modulated doubt as the criterion of mentalizing in the Reflective Functioning Questionnaire still supports the conclusion that pretend mode, psychic equivalence, and teleological mode are the main forms of non-mentalizing because

¹¹⁴ Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., ... and Lowyck, B. (2016). 'Development and Validation of a Self-Report Measure of Mentalizing: The Reflective Functioning Questionnaire'. *PLoS One*, 11(7): e0158678. For critical appraisal of the psychometric properties of the two scales of the RFQ see Spitzer, C., Zimmermann, J., Brähler, E., Euler, S., Wendt, L. P., and Müller, S. (2020). 'Die deutsche Version des Reflective Functioning Questionnaire (RFQ): Eine teststatistische Überprüfung in der Allgemeinbevölkerung'. *Psychotherapie- Psychosomatik - Medizinische Psychologie*, Early View. Müller, S., Wendt, L. P., Spitzer, C., Masuhr, O., Back, S. N., and Zimmermann, J. (2020). 'A Critical Evaluation of the Reflective Functioning Questionnaire'. <https://psyarxiv.com/5rhme/>. An earlier scale was developed by Hausberg, M. C., Schulz, H., Piegler, T., Happach, C. G., Klöpper, M., Brütt, A. L., ... and Andreas, S. (2012). 'Is a self-Rated Instrument Appropriate to Assess Mentalization in Patients with Mental Disorders? Development and First Validation of the Mentalization Questionnaire (MZQ)'. *Psychotherapy Research*, 22(6): 699–709. This scale had four factors: avoidance of thinking about mental states; difficulties perceiving inner states; psychic equivalence; and difficulties modulating affect. However, Fonagy and colleagues appear to have not known about or ignored the existence of this scale. It is not mentioned anywhere in their work, not even in the 2016 paper developing their own scale.

¹¹⁵ There is something of a contrast that can be drawn here between Fonagy and Bion in that the former emphasizes the threat of indiscriminate uncertainty. For the latter, the threat of excessive certainties is so pervasive that indiscriminate uncertainty—while likely to be avoided—is not explicitly characterized as a problem. For instance, Bion speculated that all else is ultimately disposable, but that a psychoanalytic situation is operative if neither therapist nor patient feel 'compelled to search "irritably" for certainty as a method of stifling doubts, uncertainties, mysteries, half-truths and neither is compelled to assert anything as a means by which doubt and uncertainties are evaded'. Such claims suggest that an essential strand of the psychoanalytic tradition is a commitment to generative doubt, making it seem unlikely that it would be an especially clean marker of Dynamic Interpersonal Therapy compared with other low-intensity psychodynamic therapies. Bion, W. R. ([1969] 2014). 'Further Cogitations', in Chris Mawson (ed.), *The Complete Works of W. R. Bion*, Volume 15. London: Karnac Books.. See also Ricoeur, P. (1970). *Freud and Philosophy: An Essay on Interpretation*, New Haven: Yale University Press.

¹¹⁶ It might be argued that the relationship between mentalizing and non-mentalizing would mirror adult attachment styles, which are conventionally measured on two dimensions of insecurity. However, recent psychometric work on adult attachment styles has indicated that this approach neglects a third factor, of security, distinct from the absence of insecurity. Gillath, O., Hart, J., Nofle, E. E., and Stockdale, G. D. (2009). 'Development and Validation of a State Adult Attachment Measure (SAAM)'. *Journal of Research in Personality*, 43(3): 362–373; Frias, M. T., Shaver, P. R., and Mikulincer, M. (2014). 'Measures of Adult Attachment and Related Constructs', in G. J. Boyle and D. H. Saklofske (eds), *Measures of Personality and Social Psychological Constructs*, Philadelphia, PA: Elsevier, pp. 417–447. Evidence in support of this conclusion comes from a factor analysis of aspects of a clinician-report measure of mentalizing, in which mentalizing was identified as a distinct factor, not merely the inverse of the non-mentalizing factors. Gagliardini, G. and Colli, A. (2019). 'Assessing Mentalization: Development and Preliminary Validation of the Modes of Mentalization Scale'. *Psychoanalytic Psychology*, 36(3): 249–258. It is also notable that these researchers found that MBT therapists rated 'good mentalizing' higher than CBT therapists, suggesting that they had been trained to recognize this as a distinct quality of clinical interactions.

¹¹⁷ Luyten, P., Mayes, L. C., Nijssens, L., and Fonagy, P. (2017). 'The parental reflective functioning questionnaire: development and preliminary validation'. *PLoS ONE*, 12(5): e0176218.

*the kind of mental processing they entail specifically undermines generative uncertainty and an 'inquisitive stance' towards mental states.*¹¹⁸

A further advantage of the conceptualization of mentalization in terms of the syntheses of experience and understanding is that it clarifies the status of pretend mode, psychic equivalence, and teleological mode. Fonagy has stated his suspicion that they are not 'a complete map of the nonmentalizing world'.¹¹⁹ We agree that there could be other forms of mental processing that would block generative uncertainty and access to the 'inquisitive stance'. One example has been highlighted by Winnicott: when the experiences of others are used to account for and explain the motivations and intentions of others and oneself, to the exclusion of one's own subjectively conceived and the objectively perceived experiences. And Fonagy and Target, discussing Winnicott, have indeed acknowledged this as an integral obstacle to mentalization.¹²⁰

A further candidate for consideration as a form of non-mentalizing has been raised by Ensink, Fonagy, and colleagues. They have described states of withdrawal from others as fundamentally 'incompatible' with mentalization.¹²¹ Such states could occur for various reasons—for instance, anticipated pain or anxiety associated with contact with others, whether based on their actual past behaviour or on externalization of the alien self.¹²² For instance, technology addiction has been proposed as a case in which the capacity to understand the thoughts and feelings of others is blocked by a kind of withdrawal.¹²³ Such claims are supported by empirical findings showing that extensive online game use is associated with unresolved loss and trauma on the Adult Attachment Interview, with the relationship mediated by symptoms of dissociation.¹²⁴ If Fonagy and Luyten are right that understanding one's own mind is secondary to understanding the minds of others, then sustained and intense withdrawal would prove an obstacle not just to understanding the mental states of others but also oneself.

It might be argued that withdrawal is not a distinct form of non-mentalizing, but rather a moment in the process of pretend mode, which depends upon an obstruction of sources of specific information about reality. However, if so, it is a distinct and separable moment, since not all withdrawal then prompts pretend mode. Withdrawal may even *protect* mentalizing,

¹¹⁸ By way of comparison, Luyten and colleagues drew upon a theoretical review of the concept to identify three essential elements of parental reflective function (PRF) in developing a self-report measure: '(a) interest and curiosity in mental states, (b) the ability to recognize the opacity of mental states, and (c) nonmentalizing modes characteristic of parents with (severe) impairments in PRF (e.g. malevolent attributions, inability to enter into the subjective world of the child)'. Again, the first and second factors can be regarded as reflecting a discriminated capacity for doubt and reconsideration. Luyten, P., Mayes, L. C., Nijssens, L., and Fonagy, P. (2017). 'The Parental Reflective Functioning Questionnaire: Development and Preliminary Validation'. *PLoS One*, 12(5): e0176218.

¹¹⁹ Duschinsky, R., Collver, J., and Carel, H. (2019). "'Trust Comes From a Sense of Feeling One's Self Understood by Another Mind": An Interview with Peter Fonagy'. *Psychoanalytic Psychology*, 36(3), 224–227.

¹²⁰ Target, M. and Fonagy, P. (1996). 'Playing with reality: II. The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479, p. 467. See also Fonagy, P., Moran, G. S., and Target, M. (1993). 'Aggression and the Psychological Self'. *The International Journal of Psychoanalysis*, 74: 471–485. 'Winnicott (1965) describes how a "false self", dedicated to pleasing the object, but eschewing genuine reflection, may develop' (p. 475).

¹²¹ Ensink, K., Leroux, A., Normandin, L., Biberdzic, M., and Fonagy, P. (2017). 'Assessing Reflective Parenting in Interaction with School-Aged Children'. *Journal of Personality Assessment*, 99(6): 585–595, p. 588.

¹²² See Steiner, J. (2003). *Psychic Retreats: Pathological Organizations in Psychotic, Neurotic and Borderline Patients*, London: Routledge.

¹²³ Schimmenti, A. and Caretti, V. (2010). 'Psychic Retreats or Psychic Pits?: Unbearable States of Mind and Technological Addiction'. *Psychoanalytic Psychology*, 27(2): 115.

¹²⁴ Schimmenti, A., Guglielmucci, F., Barbasio, C., and Granieri, A. (2012). 'Attachment Disorganization and Dissociation in Virtual Worlds: A Study on Problematic Internet Use among Players of Online Role Playing Games'. *Clinical Neuropsychiatry*, 9(5): 195–203.

by sustaining a degree of reserve from collective non-mentalizing cultures (see Chapter 9).¹²⁵ Nonetheless, over time, sustained withdrawal is likely to prompt pretend mode as a form of non-mentalizing. Where we withdraw 'to' is never a sphere in which others are psychologically absent. Our self-representation will nonetheless still contain narrative representations of the thoughts and feelings of others, and withdrawal will hinder the updating and modulation of these representational elements.¹²⁶

Two further, perhaps more minor forms of non-mentalisation, mentioned above, might include systematic obstacles to mentalizing thoughts or feelings. However, the capacity to mentalize thoughts but not feelings does not represent the kind of systematic obstruction to generative doubt and inquisitive stance achieved by pretend mode, psychic equivalence, or teleological mode. Yet, as Bateman and Fonagy have observed, 'with cognition there is always doubt' and therefore the potential for 'cognitive reappraisal'.¹²⁷ Likewise, strong capacities for mentalizing feelings in the context of weak capacities for mentalizing thoughts is likely to hinder the coherence and consistency of accounts and explanations of experiences and observable behaviour. However, this configuration is again unlikely to represent a systematic obstacle to mentalizing.¹²⁸

A further possibility might be difficulties mentalizing past experience, as opposed to present experience. There are innumerable ways in which present-day routines can be set up to hinder attention to thoughts and feelings from the past. Again, this would be troublesome for mentalizing, not least because comparison of past and present is a powerful aid to reconsideration. However, when consideration of thoughts and feelings in present experience is feasible, difficulties thinking about the past would rarely function as a block on mentalizing to the same extent as pretend mode, psychic equivalence, or teleological mode. These would all be minor forms of non-mentalizing, not on the same level as the pretend mode, psychic equivalence, and teleological mode for obstructing generative doubt and inquisitive stance. Yet, in the 2000s, as mentioned in Chapter 4, Fonagy, Gergely, and Target described the ability to 'represent and differentiate between the mental models of the self and of the other about reality' as one of the cognitive requisites for mentalization.¹²⁹ As such, a potent potential form of non-mentalizing, and candidate for a 'fourth' major form, is what Fonagy and colleagues call 'the externalization of the alien self'. This notion of 'externalization of the alien self', and the conceptualization of the self in general, will be the subject of the next chapter.

¹²⁵ Duschinsky, R., Reisel, D., and Nissen, M. (2018). 'Compromised, Valuable Freedom: Flat Affect and Reserve as Psychosocial Strategies'. *Journal of Psychosocial Studies*, 11(1): 68–92. See also Winnicott, D. W. (1988). 'Withdrawal and Regression', in *Human Nature*, London: Free Association Books, pp. 141–142.

¹²⁶ Winnicott, D. W. (1945). 'Primitive Emotional Development'. *The International Journal of Psychoanalysis*, 26: 137–143.

¹²⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 52.

¹²⁸ Tuber, S. (2015). 'Psychological Mindedness in the Face of a Learning Disability: The Utility of Play'. *Journal of Infant, Child, and Adolescent Psychotherapy*, 14(3): 288–293; Derks, S., van Wijngaarden, S., Wouda, M., Schuengel, C., and Sterkenburg, P. S. (2019). 'Effectiveness of the Serious Game "You & I" in Changing Mentalizing Abilities of Adults with Mild to Borderline Intellectual Disabilities: A Parallel Superiority Randomized Controlled Trial'. *Trials*, 20(1): 1–10.

¹²⁹ Fonagy, P., Gergely, G., and Target, M. (2007). 'The Parent–Infant Dyad and the Construction of the Subjective Self'. *Journal of Child Psychology and Psychiatry*, 48(3–4): 288–328, p. 290.

6

Conceptualizing the ‘self’

Introduction

In Chapter 5, we described Fonagy and colleagues’ concept of ‘hypermentalization’ as a form of non-mentalizing in pretend mode. Hypermentalization is difficult to tell apart from mentalization. The elements that seem to be absent are grounding in present or past experience and the capacity for reconsideration of thoughts and feelings. Yet the construct of perceptual experience could use further consideration, above all in how it overlaps with or differs from the concept of ‘self’. The idea of ‘the self’ has been a critical one for Fonagy and colleagues for over 30 years. The theory of the self stemming from work on mentalization has been praised by allies as representing one of the main advances over prior work on ‘theory of mind’, such as Baron-Cohen (Chapter 3).¹ Bateman and Fonagy have even described mentalizing as ‘first and foremost’ a ‘theory of the development of the self’.²

In part, this is because mentalization has developed as a clinical theory, in a context in which patients themselves have described their symptoms as relating to the ‘self’ and its problems.³ However, Fonagy and colleagues have also developed a surprising, and at times uncanny, account of the self, which has informed their theory of mentalization in diverse ways. This includes the distinction between mentalizing the self and mentalizing others, which Fonagy and Luyten gave as one of the four opposing poles of mentalization (before then qualifying that actually mentalizing self and mentalizing others is not actually an opposition). Yet, beyond this, consideration of the theory of the self is also essential for understanding how Fonagy and colleagues believe thoughts and feelings are constructed, the relationship between conscious and unconscious aspects of the self, and their account of what we glean and protect ourselves from within present or past experience.

The terms ‘mental states’ or ‘intentions’ serve as placeholders in the work of Fonagy and colleagues that help mentalization theory travel across disciplinary spaces to varied audiences, though at the price of simplification. At times, their writing can give the impression that thoughts and feelings about personal experience are lying around within minds, waiting to be recognized unless blocked by non-mentalizing.⁴ This may be an effect of Fonagy’s early indebtedness to the literature on theory of mind, in which mental states tend to be treated as singular and determinate.⁵ Fonagy and colleagues have paid little attention to differences

¹ Ensink, K. and Mayes, L. C. (2010). ‘The Development of Mentalisation in Children from a Theory of Mind Perspective’. *Psychoanalytic Inquiry*, 30(4): 301–337.

² Bateman, A. W. and Fonagy, P. (2003). ‘The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder’. *Bulletin of the Menninger Clinic*, 67: 187–211, p. 192.

³ Adshead, G. and Fonagy, P. (2012). ‘How does Psychotherapy Work? The Self and its Disorders’. *Advances in Psychiatric Treatment*, 18(4): 242–249: ‘A key reason for people seeking psychological therapy is an experience of a disordered “self”’ (p. 242).

⁴ Burman has discussed the apparent tension between the language of understanding mental states and the apparent psychoanalytic commitments of Fonagy and colleagues. Burman, E. (2016). *Deconstructing Developmental Psychology* (3rd edn), London: Routledge, p. 160.

⁵ Sedgwick, E. K. (2011). ‘Affect Theory and Theory of Mind’, in Jonathan Goldberg (ed.) *The Weather in Proust*, Durham, NC: Duke University Press.

between mentalizing unitary versus multiple or conflictual states of mind, which is surprising since their correlates could be expected to be quite different. As a point of contrast, in Bion's work, thoughts and feelings could be inchoate, hybrid, characterized by inadequacy or over-spill. He regarded this as important because whole domains of human experience are defined precisely by falling beneath or beyond determinate thoughts and feelings: this ranges from disavowed wishes, to experiences of the sublime, to taken-for-granted and sedimented habit. As Bion argued, even our most unitary and determinate thoughts and feelings are elaborated, and subjectively encountered as relevant, within a thick and heterogenous context of proto-thoughts and proto-feelings prompted by our past experiences and present environment.⁶

An interesting illustration is mood. Besides its use in the technical phrase 'mood disorders', the term appears remarkably rarely in the writings of Fonagy and colleagues, given that the theory of mentalizing is centrally concerned with mental states. This is despite the fact that they acknowledge that improvements in mood are one of the priorities they themselves have for treatment, perhaps even above symptomatic improvement.⁷ Fonagy and colleagues have, if somewhat unsteadily, generally treated mental states as if they were relatively determinate and linked with intentions. However, it has long been recognized by scholars of mood, at least since Edith Jacobson in the psychoanalytic literature, that mood does not necessarily have these qualities. Indeed, a mood does not need to be 'about' anything particular; it does not need to have a particular intended object.⁸ This is implicitly acknowledged by Fonagy and colleagues on the various occasions that the term 'mood' does appear in their writing. So, for instance, in their 1991 paper, Fonagy and colleagues state that reflective functioning developmentally incorporates skills at both 'sensing and responding to the mood of another' and 'the understanding of others' intentions', distinguishing the two.⁹ In 2008,

⁶ Bion, W. R. ([1963] 1984). *Elements of Psychoanalysis*, London: Karnac Books; Matte-Blanco, I. ([1988] 2003). *Thinking, Feeling, and Being*, London: Routledge. Converging on a similar point from within the phenomenological tradition, see Merleau-Ponty, Maurice. ([1960] 1964). *Signs*, trans. Richard C. McCleary, Evanston, IL: Northwestern University Press; Merleau-Ponty, Maurice. ([1964] 1968). *The Visible and the Invisible*, trans. Alphonso Lingis, Evanston: Northwestern University Press. According to Sperber and Wilson, the penumbra of proto-thoughts and proto-feelings that array themselves around any potential determinate thought or feeling are part of what gives the question of relevance its particular urgency. For them, this is an essential reason for the need for epistemic vigilance, in order to filter this penumbra, and to identify those elements relevant to the self and/or conversational others. This element of Sperber and Wilson's work does not appear to have been taken up by Fonagy and colleagues in their appropriation of the concept of epistemic vigilance. Sperber, D. and Wilson, D. (2004). 'Relevance Theory', in L. R. Horn and G. Ward (eds), *The Handbook of Pragmatics*, Oxford: Blackwell, pp. 607–632; de Saussure, L. and Wharton, T. (2019). 'La notion de pertinence au défi des effets émotionnels'. *Travaux Interdisciplinaires sur la Parole et le Langage (TIPA)*, 35: 1–23.

⁷ Fonagy, P. (2010). 'The Changing Shape of Clinical Practice: Driven by Science or by Pragmatics?', *Psychoanalytic Psychotherapy*, 24(1): 22–43: 'From a professional's standpoint, as from that of the ordinary member of the public, physical role limitation, physical function and pain have high priority, while those suffering disorders rate dignity and general wellbeing (mood, global assessment of life, having a partner, job, lots of social contact) as more important' (p. 34).

⁸ Jacobson, E. (1957). 'Normal and Pathological Moods: Their Nature and Functions'. *The Psychoanalytic Study of the Child*, 12(1): 73–113: 'moods seem to represent, as it were, a cross-section through the entire state of the ego, lending a particular, uniform coloring to all its manifestations for a longer or shorter period of time. Since they do not relate to a specific content or object but find expression in specific qualities attached to all feelings, thoughts and actions, they may indeed be called a barometer of the ego state ... Thus, anger at somebody or something may turn into an angry mood, love or hate into a kind or hostile mood, anxiety into an anxious mood, as soon as they have ceased to relate only to special, selected objects or notions' (pp. 75–76). See also Mulhall, S. (1996). 'Can there be an Epistemology of Moods?'. *Royal Institute of Philosophy Supplements*, 41: 191–210; Ratcliffe, M. (2010). 'The Phenomenology of Mood and the Meaning of Life', in P. Goldie (ed.), *Oxford Handbook of Philosophy of Emotion*, Oxford: Oxford University Press, pp. 349–371; Flatley, J. (2017). 'Reading for Mood'. *Representations*, 140(1): 137–158.

⁹ Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). 'The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment'. *Infant Mental Health Journal*, 12(3): 201–218, p. 204.

Baradon, Fonagy, and colleagues distinguish mental states, moods, self-esteem, and sense of social connectedness in analysing the narratives of incarcerated mothers with care for their infants.¹⁰ As we shall see in Chapter 7, factor analysis of mental health symptoms reveals mood as a distinct sub-factor with distinct correlates. Luyten and Fonagy have discussed a distinction between ‘feelings’, ‘affects’, and ‘emotions’, where affects (including mood) are characterized as a tone rather than an intentional mental state oriented towards an object. They state that all three ‘would come under this category of mentalising’.¹¹ However, as we saw in Chapter 4, the definitions of mentalizing are uneven in whether inchoate and non-intentional forms of thought and feeling form part of the construct.¹²

What account Fonagy and colleagues offer of the dynamic properties and emergence of thoughts and feelings offered is often speculative, not fully fleshed out, and discussed primarily in venues for psychoanalytic audiences.¹³ It is also generally yet to generate commentary, whereas the more basic elements of mentalization theory have generated a large literature, albeit predominantly exposition (the primary exceptions are the few dedicated critiques of the concept of implicit mentalizing).¹⁴ A contributing factor is that ideas of proto-thoughts and feelings are bound up with work by Fonagy and colleagues reflecting on the idea of the ‘self’. This is a tricky concept, with a history of contradictory use of terminology regarding the self in psychoanalytic theory especially and psychological discourse more broadly.¹⁵

Fonagy and colleagues acknowledged the problems that the term ‘self’ could cause, and took psychoanalytic theory to task for loose use of the concept, though Fonagy and colleagues not infrequently fail to heed their own warnings. Critics have stated that, given the pivotal role of the concept of ‘self’ in work on mentalization, the lack of definition of the concept is a major problem for the theory as a whole.¹⁶ Fonagy and colleagues can be found

¹⁰ Baradon, T., Fonagy, P., Bland, K., Lénárd, K., and Slead, M. (2008). ‘New Beginnings—An Experience-Based Programme Addressing the Attachment Relationship between Mothers and their Babies in Prisons’. *Journal of Child Psychotherapy*, 34(2): 240–258, p. 249.

¹¹ Luyten, P. and Fonagy, P. (2012). ‘The Multidimensional Construct of Mentalization and its Relevance to Understanding Borderline Personality Disorder’, in A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience*, Oxford: Oxford University Press, pp. 405–427, p. 412.

¹² For instance, in their 2007 definition of mentalizing (Chapter 4), Fonagy and colleagues appear to unpack emotions into ‘needs, desires, feelings, and thoughts into ‘beliefs, goals, and reasons’. The distinction here between ‘desires’ and ‘feelings’ links to long-standing psychoanalytic interest in desire as a form of intentionality (in the sense of wishes or plans regarding the past or present) that may well not be aligned with our self-representation, or readily conceived of or reconsidered. Green, A. (1999). *The Fabric of Affect in the Psychoanalytic Discourse*, New York: Routledge.

¹³ For instance, Fonagy and Allison’s theory of the nature of consciousness and of the unconscious has huge ramifications, but is put forward in a sustained way only in Fonagy, P. and Allison, E. (2016). ‘Psychic Reality and the Nature of Consciousness’. *The International Journal of Psychoanalysis*, 97(1): 5–24.

¹⁴ E.g. Kiverstein, J. (2011). ‘Social Understanding without Mentalizing’. *Philosophical Topics*, 39(1): 41–65; Davidsen, A. S. and Fosgerau, C. F. (2015). ‘Grasping the Process of Implicit Mentalization’. *Theory & Psychology*, 25(4): 434–454; Liljenfors, R. and Lundh, L. G. (2015). ‘Mentalization and Intersubjectivity towards a Theoretical Integration’. *Psychoanalytic Psychology*, 32(1): 36–60; Køster, A. (2017). ‘Mentalization, Embodiment, and Narrative: Critical Comments on the Social Ontology of Mentalization Theory’. *Theory & Psychology*, 27(4): 458–476. Another important critical commentary, addressing limitations in the acknowledgement of social context in the conceptualization of mentalizing, was Sperry, M. (2013). ‘Putting our Heads Together: Mentalizing Systems’. *Psychoanalytic Dialogues*, 23(6): 683–699. Wide swathes of the theory of Fonagy and colleagues have been left largely untouched by such commentaries, most notably the concepts of the self, primary unconscious, the different modes of non-mentalizing, and adaptation.

¹⁵ See e.g. Sandler, J. and Rosenblatt, B. (1962). ‘The Concept of the Representational World’. *The Psychoanalytic Study of the Child*, 17(1): 128–145.

¹⁶ E.g. Horne, M. (2003). ‘Affect Regulation, Mentalization, and the Development of the Self by Peter Fonagy, György Gergely, Eliot Jurist, Mary Target’. *Fort Da*, 9(2): 107–111; Ferraro, D. ([2011] 2014). ‘The Other, Clinical and Empirical: A Review of Fonagy et al. on Affect Regulation, Mentalization, and the Development of the Self’. Accessed at: <https://melbournelacanian.wordpress.com/2014/11/24/the-other-clinical-and-empirical-a-review-of-fonagy-et-al-on-affect-regulation-mentalisation-and-the-development-of-the-self>.

at times leaning on the term's imprecision for making evocative claims, in the manner they criticized in psychoanalytic discourse.¹⁷ For instance, as we saw in Chapter 3, Fonagy and colleagues have at times described disorganized attachment as caused by disorganization of the self; and at times they have described disorganization of the self as caused by disorganized attachment. Of course both may be true, but it raises the question of what is meant by 'self' in each instance. More generally, clarification of the concept of the self has important bearing for understanding the operation of both mentalizing and non-mentalizing. Additionally, recent statements have indicated that Fonagy and colleagues plan to further integrate their ideas about the self more firmly into approaches to clinical technique in the near future (see Chapter 8).

This chapter will begin by considering broader currents in social science and psychology that made the concept of 'self' salient at the point that mentalization theory was being developed, and the ways that Fonagy and colleagues have conceptualized the self. We discuss the concept of 'alien self', introduced by Fonagy and colleagues to describe the experience of desires and elements of personal experience that disturb self-representations. We will then explore the account they offer of sexuality and aggression, as two inevitable and especially potent components of the alien self.

What is a self?

Deployment of the concept of 'self' by Fonagy and colleagues can usefully be placed in the history of psychological theory. In *Principles of Psychology* in 1890, William James dedicated a chapter of 'consciousness of self', which he described as 'justly regarded as the most puzzling puzzle with which psychology has to deal; and whatever view one may espouse, one has to hold his position against heavy odds'.¹⁸ From the turn of the century, the idea of the 'self' fell into decline as part of the emergence of psychology as an empirical discipline, which distrusted the concept as a mask and obfuscation of actual psychological mechanisms that could be identified, operationalized, and tested. The primary exception was Mary Whiton Calkins, who was forced to fight the current in order to maintain experiences of the 'self' as a legitimate object of psychological theory and scientific inquiry. Calkins felt that a concept of 'self' was needed in order to understand several features of psychological life, not least how individuals experience themselves as morally or socially obligated to others.¹⁹ The current reversed, however, in the early 1950s.²⁰ Key second generation of psychoanalytic thinkers had fled Europe for New York in the 1930s and developed ego psychology, an approach that downgraded the centrality of drive theory and

¹⁷ E.g. Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 1. Borderline Personality Disorder as a Limitation of Resilience'. *Borderline Personality Disorder and Emotion Dysregulation*, 4: 11: 'It is not possible for these individuals to access mentalizing if the self is overwhelmed by negative interference.' It can be inferred here that 'self' primarily means both working memory and phenomenological experience. But the reader is left speculating about the meaning of this claim, which appears to be explanatory. In fact, the underspecified use of the concept of 'self' makes the claim descriptive and metaphoric.

¹⁸ James, W. (1890). *Principles of Psychology*, New York: Holt, p. 330. See also Baumeister, R. F. (1987). 'How the Self Became a Problem: A Psychological Review of Historical Research'. *Psychological Review*, 52(1): 163–176; Danziger, K. (1997). 'The Historical Formation of Selves', in R. Ashmore and L. Jussim (eds), *Self and Identity: Fundamental Issues*, New York: Oxford University Press, pp. 137–159.

¹⁹ Calkins, M. W. (1915). 'The Self in Scientific Psychology'. *American Journal of Psychology*, 26: 495–524.

²⁰ Hilgard, E. R. (1949). 'Human Motives and the Concept of the Self'. *American Psychologist*, 4(9): 374–382; Sarbin, T. R. (1952). 'A Preface to a Psychological Analysis of the Self'. *Psychological Review*, 59(1): 11–22.

instead accentuated the individual's adaptation to their environment.²¹ In 1950, one of the leading ego psychologists, Hartmann, introduced the term 'self' into psychoanalytic theory. He proposed that whereas 'ego' should be used for descriptions of intra-individual dynamics, the term 'self' should be used for examining the relationship of the individual with others.²² This was in part in response to an ambiguity in Freud's writings, where the term 'ego' had encompassed both senses.²³ In the same period as Hartmann's introduction of the 'self' into psychoanalysis, concepts of 'self-esteem' and 'self-actualization' were put forward by American humanistic psychologists such as Abraham Maslow and Carl Rogers.²⁴ Use of the term 'self' helped Rogers contrast his humanistic therapeutic approach with Hartmann and the position of the ego psychologists, for whom the 'ego' was to a large extent unconscious. Rogers' deployed the idea of 'self' to highlight the active capacity of patients to know their own needs and to judge the goals and appropriate termination point for therapy.²⁵

Examination of these discourses suggests that 'self' was less a determinate single object than a cluster of loosely related concerns, and their interrelation. Seven different referents of 'self' may be identified:

1. An individual as a discrete whole.
2. Her experience as an embodied subject.
3. Her passing or enduring knowledge of herself.
4. Enduring qualities of her personality, which she may or may not perceive.
5. Her agency.
6. Her social identity.
7. Her experience of personal authenticity.²⁶

²¹ Hartmann, H. (1939). *Ego Psychology and the Problem of Adaptation*, New York: International Universities Press; Jacobson, E. (1954). 'The Self and the Object World: Vicissitudes of their Infantile Cathexes and their Influence on Ideational Affective Development'. *The Psychoanalytic Study of the Child*, 9: 75–127. Though not an ego psychologist, Karen Horney was also important in the development of 'self' discourse in psychoanalysis, again as a reaction to drive theory. Horney, K. (1950). *Neurosis and Human Growth*, New York: W. W. Norton.

²² Hartmann, H. ([1950] 1964). 'Comments on the Psychoanalytic Theory of the Ego', in *Essay on Ego Psychology*, New York: International Universities Press, pp. 113–141.

²³ Fromm, E. (1970). 'Freud's Model of Man and its Social Determinants', in E. Fromm (ed.), *The Crisis of Psychoanalysis: Essays on Freud, Marx, and Social Psychology*, New York: Holt, Rinehart, Winston, pp. 42–61; Bettelheim, B. (1984). *Freud and Man's Soul*, New York: Vintage; Borch-Jacobsen, M. (1988). *The Freudian Subject*, Stanford: Stanford University Press.

²⁴ Maslow, A. H. (1942). 'Self-esteem (dominance-feeling) and sexuality in women'. *The Journal of Social Psychology*, 16(2): 259–294; Maslow, A. H. (1948). "Higher" and "Lower" Needs'. *The Journal of Psychology*, 25(2): 433–436; Rogers, C. R. (1950). 'The Significance of the Self-Regarding Attitudes and Perceptions', in M. L. Reymert (ed.), *Feelings and Emotions: The Mooseheart Symposium*, New York: McGraw-Hill. Rogers, C. and Dymond, R. (1954). *Psychotherapy and Personality Change*, Chicago: University of Chicago Press. The term 'self' also came to be used by the ego psychologists.

²⁵ Rogers, Carl. (1942). *Counseling and Psychotherapy: Newer Concepts in Practice*, Boston; New York: Houghton Mifflin Company; Rogers, C. R. (1951). *Client-Centered Therapy: Its Current Practice, Implications, and Theory*, Boston: Houghton Mifflin Company; Symonds, P. M. (1951). *The Ego and the Self*, New York: Appleton-Century-Crofts.

²⁶ Our synthesis of seven referents in self discourse since the 1950s overlaps for the most part with Leary and Tangney's distinctions between five meanings of the term 'self' in psychology: self as total person; self as personality; self as experiencing subject; self as beliefs about oneself; and self as executive agent. However, the importance of the use of 'self' to mean experience of personal authenticity was central to its use in the 1950s, and is also very relevant to Fonagy and colleagues—as was the use of 'self' to mean social identity in discussions about the receptivity of individuals to social and developmental influence. Leary, M. R. and Tangney, J. P. (2003). 'The Self as an Organizing Construct in the Behavioral and Social Sciences', in M. R. Leary and J. P. Tangney (eds), *Handbook of Self and Identity*, New York: Guilford, pp. 3–14.

An important commentator at the time, Gordon Allport, called use of the term self 'lazy'. He felt that the apparent obviousness and utility of the term lured psychologists into failing to draw distinctions between its varied meanings.²⁷ It was easy for psychologists to talk right past one another, or their audiences, simply by assuming different meanings of the term. Allport felt that the concept appealed in part because of clear connections between these different phenomena, but simultaneously served to obscure their precise nature. Allport also astutely highlighted the social context of attention to the 'self', reflecting on the growing importance of individual self-management and personal lifestyle in the post-war economy and culture.²⁸ Individual self-management and personal lifestyle would only grow as objects of concern and discussion over the subsequent decades (see Chapter 2). In the 1960s and 1970s, the 'self' would be re-absorbed by psychoanalysis in the work of Winnicott and Kohut. These psychoanalytic thinkers argued that failures of parental empathy and soothing responsiveness would lead developmentally to i) inauthentic and inaccurate self-knowledge on the part of individuals, ii) weaknesses and incoherencies in personality structure, iii) challenges in modulating emotions, and iv) difficulties with social relationships. The reason lay, Winnicott and Kohut argued, in the fact that all four are formed on the bedrock of early child-caregiver relationships.²⁹

Commenting on these developments in their early book, *Personality Theory and Clinical Practice*, published in 1985, Fonagy and Higgitt criticized Rogers and humanistic psychologists for treating 'the self as an entity and a causal agent able to consider and integrate perceptions'.³⁰ This results in circular explanations, in which the role of thoughts and feelings get missed because motivation is simply assigned to 'the self' as the cause of action or beliefs. Fonagy was more sympathetic to the position of Winnicott and Kohut, whom he regarded as invoking the concept of 'self' in analysing the developmental outcomes of early care.³¹ In the 1990s, Fonagy and Target's own emerging theory of mentalization aligned well, and to an extent was influenced by Winnicott and Kohut's emphasis on the caregiver's attention to the child's thoughts, feelings, and/or intentions. Like these earlier thinkers, Fonagy and Target regarded qualities of the caregiver's attention to the child as

²⁷ Allport, G. W. (1955). *Becoming; Basic Considerations for a Psychology of Personality*, Volume 20, London: Yale University Press, p. 38.

²⁸ See also Cushman, P. (1995). *Constructing the Self, Constructing America: A Cultural History of Psychotherapy*, Cambridge, MA: Perseus; Shaw, R. and Colimore, K. (1988). 'Humanistic Psychology as Ideology: An Analysis of Maslow's Contradictions'. *Journal of Humanistic Psychology*, 28: 51–74; Carrette, J. (2003). 'Psychology, Spirituality and Capitalism: The Case of Abraham Maslow'. *Critical Psychology*, 8: 73–95; Malone, K. R. (2007). 'The Subject as Drop-Out: Cultural Accountability and the Ethics of Psychoanalysis and Humanistic Psychology'. *Theory & Psychology*, 17: 449–471.

²⁹ Winnicott, D. W. ([1960] 1965). *Ego Distortion in Terms of True and False Self, the Maturation Process and the Facilitating Environment*, New York: International Universities Press, pp. 140–152; Kohut, H. (1971). *The Analysis of the Self*, New York: International Universities Press; Kohut, H. (1977). *The Restoration of the Self*, New York: International Universities Press. See also Harwood, I. (1987). 'The Evolution of the Self: An Integration of Winnicott's and Kohut's Concepts', in T. Honess and K. Yardley (eds), *Self and Identity: Individual Change and Development*, London: Routledge, pp. 55–77; Lunbeck, E. (2014). *The Americanization of Narcissism*, Cambridge, MA: Harvard University Press.

³⁰ Fonagy, P. and Higgitt, A. C. (1985). *Personality Theory and Clinical Practice*, London: Methuen, p. 92.

³¹ Bach, S., Mayes, L., Alvarez, A. and Fonagy, F. (2000). 'Panel 1: Definition of the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24. Fonagy: 'It was from instances of its apparent falsity or damage that psychoanalysis first sought to define what a coherent self might be. Consider, for example, Deutsch (1942) and the "as if" personality, Winnicott (1965) and the false self, or Kohut (1971) and the grandiose self... Paradoxically, when selfhood is authentic there is little need for the concept. The individual recognizes both impulse and prohibition; the compromise between them forms an attribute of the ego. It is when this process fails that the critical role of the self-structure asserts itself most urgently' (pp. 20–1).

consequential for their later self-understanding, personality, capacities for affect regulation, and social relationships.³²

Such an account also aligned with the conclusions of attachment theory in the late 1980s and early 1990s—for instance, in the work of Alan Sroufe³³ (see Chapter 3). These different perspectives all converged on an account in which the child's sense of self is 'originally an extension of experience of the other'.³⁴ This position became an abiding commitment for Fonagy from the 1990s, and one that he repeatedly identified with Hegel's claim that 'self consciousness exists in and for itself when, and by the fact that, it so exists for another; that is, it exists only in being acknowledged'.³⁵ The passage from Hegel was interpreted by Fonagy and colleagues to mean that the child's capacity to conceive of thoughts and feelings within itself stems from the reflective functioning of the caregiver, who provides emotional containment and acknowledgement of the child's mental states and/or intentions.³⁶ At times, Fonagy and colleagues extended this claim, to argue that the capacity to mentalize the self in adulthood or even the experience of having a self is an extension of interaction with and acknowledgement of others.³⁷ However, this has been an unsteady claim. As we saw in Chapter 4, Bateman, Fonagy, and Campbell have recently claimed the opposite, that 'to mentalise others requires the capacity to mentalise the self'.³⁸ Examining the definitions of mentalizing in the final section of Chapter 4, it can be seen that sometimes Fonagy and colleagues imply that we understand the thoughts and feelings of

³² See the discussions of Winnicott and Kohut in Target, M. and Fonagy, P. (1996). 'Playing with Reality: II. The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479.

³³ Sroufe, L. A. (1989). 'Relationships, Self, and Individual Adaptation', in A. J. Sameroff and R. N. Emde (eds), *Relationship Disturbances in Early Childhood: A Developmental Approach*, New York: Basic Books, pp. 70–94: 'Self should be conceived as an inner organisation of attitudes, feelings, expectations, and meanings, which arises from an organized caregiving matrix' (p. 71).

³⁴ Target, M. and Fonagy, P. (1996). 'Playing with Reality: II. The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479, p. 474.

³⁵ Hegel, G. W. F. ([1807] 1977), *Phenomenology of Spirit*, trans. A. V. Miller, ed. J. N. Findlay, Oxford: Oxford University Press, p. 111. The first published citation of this passage by Fonagy was in Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration'. *Bulletin of the Anna Freud Centre*, 17: 33–55; and then next in Fonagy, I. and Fonagy, P. (1995). 'Communication with Pretend Actions in Language, Literature and Psychoanalysis'. *Psychoanalysis and Contemporary Thought*, 18(3): 363–418. However, given how protracted the development of the latter paper was, there is reason to suspect that it is in fact the earlier work.

³⁶ Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Matton, G., and Target, M. (1995). 'Attachment, the Reflective Self and Borderline States: The Predictive Specificity of the Adult Attachment Interview and Pathological Emotional Development', in S. Goldberg, R. Muir, and J. Kerr (eds), *Attachment Theory: Social, Developmental, and Clinical Perspectives*, New York: Analytic Press, pp. 233–278, p. 256. A sharp contrast with traditional psychoanalytic theory was drawn in Target, M. and Fonagy, P. (1996). 'Playing with Reality: II: The Development of Psychic Reality from a Theoretical Perspective'. *The International Journal of Psychoanalysis*, 77: 459–479: 'For Freud, in infancy and early childhood, others in the external world were extensions of the self. While this may be an accurate description of the phenomenology, for us it seems more accurate to see the self as originally an extension of experience of the other' (p. 474).

³⁷ E.g. Luyten, P. and Fonagy, P. (2012). 'The Multidimensional Construct of Mentalization and its Relevance to Understanding Borderline Personality Disorder', in A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience*, Oxford: Oxford University Press, pp. 405–427: 'The theoretical claim that self and other representations are shared and are the default mode of the motor system is reasonably well supported, as we have seen. People have a tendency to mirror actions automatically ... the inhibition of imitative behaviour involves cortical areas that are also related to mentalizing ... the capacity to inhibit imitative behaviour may be key to enabling us to generate a sense of "me"-ness through achieving a "not other"ness. In other words, each time we interpret the actions of another, there may be a sequence in which an initial imitative matching response with the other within a motor neuron self-other system interacts with the reflective mentalizing self-other system ... This might explain why patients with BPD [borderline personality disorder] feel vulnerable to losing a sense of self' (pp. 415–416).

³⁸ Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Antisocial Personality Disorder in Community and Prison Settings', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 335–349, p. 347.

ourselves through much the same means as we do the thoughts and feelings of others. And at other times they imply that, whereas the mental states of oneself require objectification and interpretation ('seeing ourselves from the outside'), the mental states of others require inference and interpretation ('seeing others from the inside'). When this contrast is drawn, the self's mental states need to be 'recognized', whereas the mental states of others need to be 'attributed'.³⁹

From the 1990s to the present, the question of whether adults know the minds of others in the same way as we know our own minds has dogged Fonagy and colleagues as an unresolved issue.⁴⁰ However, they have been entirely consistent on the developmental claim that a child's self-representation is facilitated by parental reflective functioning. When this occurs, not only will the child's capacities to form representations of thoughts and feelings be more elaborate and sophisticated, Fonagy and colleagues anticipated that representations of the child's thoughts and feelings would also be comparatively more tolerably authentic and accurate. In the context of parental reflective functioning, Fonagy argued in 2000 that 'the child's emerging self-representation will map on to what could be called a primary or constitutional self (the child's experience of an actual state of being, the self as it is)', their perceptual experience.⁴¹ This growing focus on the importance of self-representation led Fonagy and Target to ambivalence about the way that Kohut especially, and Winnicott to an extent, invoked the idea of 'self' in explaining psychological symptoms. In 2003, Fonagy and Target argued that 'the self is presented by Kohut in representational terms, yet he ascribes motivational properties to it. In this way, the self denotes most, if not all, of the personality and therefore becomes a superfluous term, much as the concept of ego was over-extended by ego psychologists.' As a consequence, 'by fitting all psychopathology into self-defects, Kohut has homogenised psychological disorder too much'.⁴²

Fonagy agreed with previous commentators that psychoanalytic references to the 'self' had inherited from Freud's ego the capacity to refer 'alternately the individual, one part of a psychic structure, or the experiencing, subjective self'.⁴³ In addition, however, he felt that the term had gained the capacity to refer to an individual's self-concept, and to their motivations. To call all these 'the self', as did Kohut and sometimes Winnicott, was no doubt an evocative, metaphorical usage. But it was also a recipe for complacency, as a metaphorical description was substituted for work to identify precise explanation: 'Confusion and ambiguity surround the status of the concept of self in psychoanalysis, particularly whether it is a theoretical (meta-psychological) construct of a system in the mind, part of our psychoanalytic theory of the mind, or an experiential one relating to the person's cumulative affective experience'.⁴⁴ Fonagy and Target's

³⁹ Bach, S., Mayes, L., Alvarez, A. and Fonagy, F. (2000). 'Panel 1: Definition of the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 19.

⁴⁰ Køster, A. (2017). 'Mentalization, Embodiment, and Narrative: Critical Comments on the Social Ontology of Mentalization Theory'. *Theory & Psychology*, 27(4): 458–476.

⁴¹ Fonagy, P. (2000). 'Attachment and Borderline Personality Disorder'. *Journal of the American Psychoanalytic Association*, 48(4): 1129–1146, p. 1136. Earlier, immediate personal experience had been characterized as the 'prereflective self' by Fonagy and colleagues. Fonagy, P., Steele, M., Steele, H., Moran, G. S., and Higgitt, A. C. (1991). 'The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and its Significance for Security of Attachment'. *Infant Mental Health Journal*, 12(3): 201–218, p. 203.

⁴² Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 183.

⁴³ Bach, S., Mayes, L., Alvarez, A., and Fonagy, F. (2000). 'Panel 1: Definition of the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 17.

⁴⁴ Fonagy, P. (1995). '2: Peter Fonagy'. *British Journal of Psychotherapy*, 11(4): 575–584, p. 575.

criticism of Kohut and other psychoanalytic theorists would later be amplified by Luyten and Fonagy: 'These metaphors are tremendously helpful from a phenomenological perspective but also have led to the reification of these self-experiences, as if we truly "have" a false or fragile self, or that we "have" an ideal and an actual self. Although helpful clinically, they provide a metaphorical description of the phenomenological experience of depression, rather than a true explanation.'⁴⁵ The 'self' is a term for psychological theorists to conjure with, offering a luminous feeling of comprehension. The audience tend to nod along, given that there is usually some among the varied meanings of the term that both resonate and feel urgent for each person. Indeed, if psychic equivalence means to account for and explain thoughts and feelings and observable social behaviour in terms of immediate experience (see Chapter 5), then appeal to a hazy concept of 'self' by psychological theorists essentially erects a screen for diverse projections of immediate experience, and a lure towards this form of non-mentalizing.

Fonagy and Target were impressed with the use of the term 'self' by Edith Jacobson and Joseph Sandler, who had both restricted the word to refer to an individual's representations of their embodied experience and personality, and avoided using the term to refer to the totality of the person. For Jacobson, for example, the term referred to the individual's representation of the bodily and mental qualities of the individual.⁴⁶ Sandler took this further, generally avoiding the unqualified term 'self' except where meaning would be clear from context, and preferring the term 'self-representation'.⁴⁷ In his later work, Sandler amended his formulation to clarify that what is represented in the self-representation is always the self-in-interaction-with-others rather than any asocial essence to the individual.⁴⁸ Fonagy appreciated the way that Jacobson and Sandler were careful in their uses of the concept, and he agreed that any notion of the self is in part a representation. However, he personally disagreed with quite such a radical curtailment of the term. In 2000, he offered 'a definition of the self as a part of the mind that is capable equally of recognizing its mental activities as its own and of attributing mental states to others. It is more than a representation, because it has the capacity to shape and determine how representations (of itself, of others) will be formed.'⁴⁹ For Fonagy, the self is a special kind of representation that is implicated in the conception of mental states.

⁴⁵ Luyten, P. and Fonagy, P. (2016). 'The Self in Depression', in M. Kyrios, R. Moulding, M. Nedeljkovic, S. S. Bhar, G. Doron, and M. Mikulincer (eds), *The Self in Understanding and Treating Psychological Disorders*, Cambridge: Cambridge University Press, pp. 71–81, p. 72. Cf. Fonagy, P. (1999). 'Relation of Theory and Practice in Psychodynamic Therapy'. *Journal of Clinical Child Psychology*, 28(4): 513–520: 'Psychodynamic clinicians found a way around the empirical problems created by partially incompatible formulations that nevertheless needed to be employed concurrently. They loosened the definition of all the categories under consideration' (p. 518).

⁴⁶ Jacobson, E. (1954). 'On Psychotic Identifications'. *The International Journal of Psychoanalysis*, 35: 102–108, p. 102. See also Jacobson, E. (1964). *The Self and the Object World*, New York: International Universities Press. Fonagy, P. (1995). '2: Peter Fonagy'. *British Journal of Psychotherapy*, 11(4): 575–584: 'Jacobson (1964) and Schafer (1968) both made, to me, helpful distinctions between self as the totality of the person and self-representation' (p. 575).

⁴⁷ Sandler, J. and Rosenblatt, B. (1962). 'The Concept of the Representational World'. *The Psychoanalytic Study of the Child*, 17(1): 128–145: 'By the self-representation we mean that organization which represents the person as he has consciously and unconsciously perceived himself, and which forms an integral part of the representational world' (p. 134). See also Fonagy, P. (2005). 'An Overview of Joseph Sandler's Key Contributions to Theoretical and Clinical Psychoanalysis'. *Psychoanalytic Inquiry*, 25(2): 120–147.

⁴⁸ Sandler, J. (1993). 'On Communication from Patient to Analyst: Not Everything is Projective Identification'. *The International Journal of Psychoanalysis*, 74: 1097–1107.

⁴⁹ Bach, S., Mayes, L., Alvarez, A., and Fonagy, F. (2000). 'Panel 1: Definition of the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 19.

The primary unconscious

The argument of Fonagy and colleagues about the relationship between mentalization and the self as a special kind of representation is intriguing and complex. Its basis lies in their account of perceptual experience, and its psychodynamic underpinnings. For Hegel and the phenomenological tradition, human experience emerges out of the prompting of the outside world, and an individual's perceptual transaction with these prompts.⁵⁰ It is out of this openness that self-perception, and the potential to doubt our perceptions, emerges. Fonagy and colleagues generally agreed with this premise, though they have framed it in their own ways. First, as we have seen (Chapter 3), Fonagy bracketed the cultural constitution of the subject, focusing very largely on the parent-child relationship as the pivotal 'outside' that prompts the emergence of the individual subject. He would return, however, to the question of wider cultural factors in recent years (Chapter 9). Second, for Fonagy and colleagues, perceptual experience is staged in a mental setting already shaped and populated by unconscious processes.

In a paper from 2016, Fonagy and Allison have distinguished three forms of unconscious process. First are mental processes that are simply not conscious, like the breathing reflex. These can become the object of consciousness and can be controlled, but are not in themselves potentially disruptive of the self-representation and its activity. From André Green's reading of Hegel and Freud, Fonagy and Allison draw a second form of unconscious process: mental states of fragmentation, incoherence, and meaninglessness.⁵¹ These states are not with reference to, or 'about', anything particular (Brentano's sense of 'intention'); nor do they constitute motivations regarding wishes or plans for the present or past (Dennett's sense of 'intention').⁵² Fonagy and Allison termed this the 'primary unconscious'.⁵³ The primary unconscious forms an irreducible backdrop to perceptual experience of the world.

For Green, states of fragmentation, incoherence, and meaninglessness are essential to the creation of meaning. They contour the edges of any meaning, and keep away other potential meanings from the one found to be relevant. Any questioning or vigilance regarding a meaning makes use of these states, where they can serve as erasers, scrubbing out or permitting alteration of existing beliefs, and allowing for detachment from existing commitments, values, and desires.⁵⁴ They are part of the mind's 'unbinding' function. We need to have a sense of what would be nonsense in order to distinguish what is relevant to us and viable as a meaning, and to revise this sense in light of circumstance of new information. Less

⁵⁰ Ricoeur, P. (1994). *Oneself as Another*, Chicago, IL: University of Chicago Press.

⁵¹ The influence of Jim Grotstein's work on Fonagy and Allison is difficult to pick apart. On the one hand, there are few elements in Fonagy and Allison's account of the primary unconscious that are not already present in Grotstein's earlier work. However, Grotstein's account is awash with different ideas, which are not especially integrated. Fonagy and Allison cite Grotstein, but do not discuss his work. A reason may be that to do so would have required quite an extensive and subtle discussion; it may have been more economical to just include the citation. Nonetheless, there are points in Grotstein that could have perhaps benefited Fonagy and Allison's model, such as the distinction between nothingness and meaninglessness. Grotstein, J. S. (1990). 'Nothingness, Meaninglessness, Chaos, and the "Black Hole" I—The Importance of Nothingness, Meaninglessness, and Chaos in Psychoanalysis'. *Contemporary Psychoanalysis*, 26: 257–290.

⁵² Fonagy, P. and Allison, E. (2016). 'Psychic Reality and the Nature of Consciousness'. *The International Journal of Psychoanalysis*, 97(1): 5–24: 'As we move towards the primary unconscious, the mental world loses its 'aboutness', its intentional character, the quality rooted in the dyadic consciousness of marked contingent mirroring or even its failed derivatives through projective identification' (p. 14).

⁵³ Ibid.

⁵⁴ Green, A. ([1993] 1999). *The Work of the Negative*, trans. A. Weller, London: Free Association Books; Green, A. (1998). 'The Primordial Mind and the Work of the Negative'. *The International Journal of Psychoanalysis*, 79: 649–665.

positively, when individuals have had experiences that lead them to fear states of fragmentation, incoherence, and meaninglessness, then this can hold them in place, making change and learning from experience more difficult, and contributing to credulity. Fonagy and Allison, however, do not seem sure whether they would assign the primary unconscious the 'unbinding' role it is assigned by Green.⁵⁵ And they do not return to the idea of the primary unconscious, or to Green's work, in their subsequent discussions of epistemic vigilance and credulity. In their paper, the emphasis is rather on the way that the primary unconscious enters into determinate mental states as breakdowns, disruptions, or dislocations.⁵⁶ In characterizing states of fragmentation, incoherence, and meaninglessness, Fonagy and Allison are explicit that these mental states are not marked by particular intentions. Indeed, when clinicians mischaracterize such disruptions experienced by patients in terms of motivations and intentions—i.e. mentalize them—this tends to exacerbate the states of fragmentation, incoherence, and meaninglessness. By contrast, Allison and Fonagy reported that supportive interventions that do not treat such states as representing intentions tend to be more clinically effective.

Distinct from the primary unconscious, what Fonagy and Allison have described as 'the psychoanalytic unconscious' is formed by disturbing or disruptive 'intentions' (in the sense of wishes and plans) with reference to the self, others, or the world. The idea is a kind of re-description of Freud's description of the unconscious in terms of wishes as the origin of the conflicts and slips of everyday life, or Klein's description of the unconscious in terms of phantasies invested in day-to-day perceptions and interactions.⁵⁷ For Fonagy, 'Freud's major discovery, is that what the philosophers of mind call the intentional stance, which Richard Hopkins calls the sentential stance, applies to non-conscious as well as conscious mental functioning.'⁵⁸ This account of Fonagy and Allison of wishes as intentional may be regarded as a revision of Freud's original concept of libido as essentially object-seeking (i.e. pertaining to wishes or plans regarding the past or present), a model set out most influentially in *The Unconscious* in 1915.⁵⁹ The idea that intentionality characterizes non-conscious as well as conscious mental states is a distinct, intriguing, and arguable claim about psychological life. Certainly there are other theorists, including Freud in his later work, who have discussed non-conscious mental states that lack intentionality, most importantly some

⁵⁵ Relevant here perhaps is Winnicott's distinction between unintegration (relaxation of an integration) and disintegration (the breakdown of an integration), a distinction that Green tends to elide. Winnicott, D. ([1962] 1965). 'Ego Integration in Child Development', in *The Maturation Processes and the Facilitating Environment*, New York: International Universities Press, pp. 56–63; Winnicott, D. W. (1988) 'Chaos', in *Human Nature*, London: Free Association Books, pp. 135–138.

⁵⁶ However, elsewhere Fonagy and colleagues have considered meaninglessness in ways that seem to correspond to Green's characterization e.g. Fonagy, P. (2011). 'Discussion of Juan Pablo Jimenez's Paper, "A Fundamental Dilemma of Psychoanalytic Technique. Reflections on the Analysis of a Perverse Paranoid Patient"', in J. P. Jimenez and R. Muguillansky (eds), *Clinical and Theoretical Aspects of Perversion*, London: Karnac Books, pp. 63–76: 'The hallmark of the pretend mode is the experience of meaninglessness in the midst of the act of symbol creation' (p. 69), though it may be that for Fonagy and colleagues there is more than one kind of experience of meaninglessness, and only sometimes is this inflected by the primary unconscious.

⁵⁷ Fonagy, P. and Allison, E. (2016). 'Psychic Reality and the Nature of Consciousness'. *The International Journal of Psychoanalysis*, 97(1): 5–24: 'The implicit great threat posed by the primary unconscious is to the intentional quality or aboutness of experience—the feeling that it is being directed toward something. We take the view that the primary unconscious is not object related, and here we diverge from the Kleinian position that phantasies, that is, representations of instinctual aims towards objects, are the 'primary content of unconscious mental processes' (Isaacs, 1948, p. 81). The intrusions of the primary unconscious undermine the implicitly purposeful character of human experience, replacing intentionality with diffused meaninglessness' (p. 15).

⁵⁸ Low, J. (2003). 'Psychoanalysis—It is a Signifier'. *Journal of the Centre for Freudian Analysis and Research*, 13, citing from an interview with Peter Fonagy.

⁵⁹ Freud, S. ([1915] 2001). *The Unconscious. Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 14, London: Vintage.

forms of anxiety.⁶⁰ The claim that intentionality is an inherent characteristic of the psychoanalytic unconscious would seem to suggest a relationship with teleological mode. However, these links are not drawn by Fonagy and colleagues, perhaps because the concepts have been developed for somewhat different audiences.

A valuable point of comparison for Fonagy and Allison's overall account of the unconscious is Ronald Fairbairn.⁶¹ Fairbairn likewise divides the unconscious between states of fragmentation and aspects of the self-representation. However, there are critical differences. Fonagy and Allison offer no account of the origin or developmental function of the primary unconscious. Fairbairn presents an account in terms of a child's feelings of not being loved, which affect all of us to varying degrees according to his argument. Fonagy and Allison do not discuss the interaction between the primary and psychoanalytic unconscious. By contrast, Fairbairn theorizes that the psychoanalytic unconscious is elaborated as a response to the feelings of fragmentation, incoherence, and meaninglessness that stem from feeling unloved. Finally, Fonagy and Allison appear to populate the psychoanalytic unconscious only with disturbing or disruptive intentions. By contrast, Fairbairn's unconscious is a much more diverse and sociable place, including, additionally, internalized representations of loved ones with whom we sustain unconscious and preconscious inner dialogue.

Allison and Fonagy are in agreement with Fairbairn, however, in their argument that an important developmental foundation of the psychoanalytic unconscious lies in perceived failures or disruptions of acknowledgement and acceptance by caregivers.⁶² This includes many wishes and plans that are 'nasty and horrid'; Fonagy and colleagues hold that the lack of social acknowledgement and acceptance has good reason!⁶³ However, avoiding the determinism of Fonagy's earlier discussions of the role of infancy, Allison and Fonagy argued that the psychoanalytic unconscious can be elaborated throughout life. Significant or sustained failures of interpersonal acknowledgement in the context of high arousal can feed the psychoanalytic unconscious. Traumatic experiences, in particular, add to the psychoanalytic unconscious by offering a part-mirror to states of fragmentation, incoherence, and meaninglessness from the primary unconscious.⁶⁴ As a result, these become fastened to certain

⁶⁰ Freud, S. ([1926] 2001). *Inhibitions, Symptoms and Anxiety*. *Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 22, London: Vintage, pp. 75–174; Arbiser, S. and Schneider, J. (eds) (2018). *On Freud's Inhibitions, Symptoms and Anxiety*, London: Routledge. For discussions of the concepts of affect and intention in Freud, as well as their limitations, see Green, A. (1999). *The Fabric of Affect in the Psychoanalytic Discourse*, New York: Psychology Press; Johnston, A. (2010). 'Affekt, Gefühl, Empfindung: Rereading Freud on the Question of Unconscious Affects'. *Qui Parle*, 18(2): 249–289. It is not clear whether all non-conscious mental states that, by degrees, lack intentionality would be assimilated by Fonagy and colleagues to the primary unconscious. Some, such as Kierkegaard, have argued that apparently non-intentional affects, such as anxiety, may seem to intend no specific object because they actually refer to our relationship with the world as a whole. See e.g. Kierkegaard, S. ([1844] 2013). *Concept of Anxiety: A Simple Psychologically Orienting Deliberation on the Dogmatic Issue of Hereditary Sin*. *Kierkegaard's Writings*, Volume 8, Princeton, NJ: Princeton University Press. Other affects may have a similar structure, such as some forms of shame.

⁶¹ Fairbairn, W. R. D. (1943). 'The Repression and the Return of Bad Objects (with Special Reference to the "War Neuroses")'. *British Journal of Medical Psychology*, 19(3–4): 327–341; Fairbairn, W. R. D. (1944). 'Endopsychic Structure Considered in Terms of Object-Relationships'. *The International Journal of Psychoanalysis*, 25: 70–92.

⁶² Fonagy, P. and Target, M. (1998). 'Mentalization and the Changing Aims of Child Psychoanalysis'. *Psychoanalytic Dialogues*, 8(1): 87–114: 'We suggest that developmental personality disturbances arise first from the child's failure to find the image of his mind, his experience of himself as a thinker of thoughts, believer of ideas, feeler of emotions, in the mind of the caregiver (see Fairbairn, 1952)' (p. 93).

⁶³ Duschinsky, R., Collver, J. and Carel, H. (2019). '"Trust Comes from a Sense of Feeling One's Self Understood by Another Mind": An Interview with Peter Fonagy'. *Psychoanalytic Psychology*, 36(3): 224–227.

⁶⁴ Fonagy, P. and Allison, E. (2016). 'Psychic Reality and the Nature of Consciousness'. *The International Journal of Psychoanalysis*, 97(1): 5–24: 'Neglect, physical abuse, emotional maltreatment, and all forms of adversity mirror and thereby, in our model, render potentially accessible to consciousness states of mind that would ('within an average expectable environment'—Hartmann, 1950) remain far from conscious subjectivity in the primary unconscious. Neglectful, aggressive or sexually seductive parenting part-mirrors the states of destructiveness, isolation

memory content and enter the psychoanalytic unconscious in the form of intrusive thoughts or feelings lacking an immediate object.⁶⁵ Fonagy, Allison, and Campbell have also recently hinted that experiences that prompt pervasive distrust in the claims of others as dependable, generalizable, or relevant (see Chapter 7) can also contribute to states of ‘meaninglessness’, implicating a role for the primary unconscious.⁶⁶

Fortunately, most of the time, unconscious intentions or negative states do not surface into the conscious meanings that organize thoughts and feelings. However, they certainly can and often do influence the wishes, fears, and inchoate senses that inflect and give colour and tone to our primary meanings (see Chapter 1). For Fonagy and colleagues, the idea of determinate thoughts and feelings belonging to the self, or thoughts and feelings belonging to another, is therefore a post-hoc description. Seen in context, this was not an unprecedented claim among psychoanalytic discussions of the ‘self’. Fonagy cited William Grossmann who had argued that:

The ‘self’ appears to be both supremely subjective yet also an objective organization, an organismic property, discernible by others. This apparent objectivity of ‘the self’ arises from the fact that a person and those around him may equate observable and characteristic behavioral organizations or traits with an internal entity, ‘the self’. In calling the self-concept a theory or a fantasy, I do not wish in any way to diminish its importance in regulating behavior. It seems to me that . . . it organizes and directs behavior.⁶⁷

From 2000 onwards, Fonagy claimed an aligned position. For him, ‘the illusion of identity is adaptive because predicting behavior of others, as well as our own behavior, is significantly simplified by the attribution of mental-state motives. In order to be able to predict, we must assume consistency amongst these mental states. Thus, underlying the intentional stance is an idealization of the self, as well as of the self of the other.’⁶⁸ It is an idealization that permits us to generate simplified and sanitized thoughts and feelings from past experience, and to use these to account for our embodied, perceptual experiences and behaviour. For instance, Bateman and Fonagy define ‘feelings’ as ‘the conscious experience of the body state during

and despair that are perhaps ubiquitous, if occasional states of mind. When the child’s environment is contingent with (matches) such devastating mind states, the part-mirroring will bring these negative states of mind closer to subjectivity’ (p. 15).

⁶⁵ Fonagy, P. and Allison, E. (2015). ‘A Scientific Theory of Homosexuality for Psychoanalysis’, in A. Lemma and P. E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, Hove, UK: Routledge, pp. 125–137, p. 133.

⁶⁶ Fonagy, P., Allison, E., and Campbell, C. (2019). ‘Mentalising, Resilience and Epistemic Trust’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 73.

⁶⁷ Grossman, W. I. (1982). ‘The Self as Fantasy: Fantasy as Theory’. *Journal of the American Psychoanalytic Association*, 30(4): 919–937, p. 926. Fonagy cites the conference presentation, on which the 1982 paper was based, in his remarks in Bach, S., Mayes, L., Alvarez, A., and Fonagy, F. (2000). ‘Panel 1: Definition of the Self’. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 17, citing Ticho, E. A. and Richards, A. D. (1982). ‘Psychoanalytic Theories of the Self’. *Journal of the American Psychoanalytic Association*, 30(3): 717–733. He also cites a review of Grossman and Kohut’s concepts of the self: Havens, L. (1986). ‘A Theoretical Basis for the Concepts of Self and Authentic Self’. *Journal of the American Psychoanalytic Association*, 34(2): 363–378. Fonagy may have additionally been influenced by Laplanche’s interpretation of the ego in Freud as a fantasy. Laplanche, J. (1976). *Life and Death in Psychoanalysis*, Baltimore, MD: John Hopkins University Press. Fonagy would also have been doing research towards his book with Hepworth on psychoanalytic theories, which mentions in passing Lacan’s characterization of the ego. For Lacan, at least in his early work, the ego is as an imaginary function, one that operates within psychological life only as a representation. Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 17.

⁶⁸ Bach, S., Mayes, L., Alvarez, A., and Fonagy, F. (2000). ‘Panel 1: Definition of the Self’. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 22.

emotional activation', acknowledging that this is a mediated and partial experience incorporating but irreducible to the body's inchoate and heterogeneous reactions and states.⁶⁹ The exact similarities and differences between the formation of feelings and thoughts are not at all well spelt out by Fonagy and colleagues. Appeal to the omnibus term 'mental states' has rather masked this question. Nonetheless, Fonagy and colleagues give the impression that a similar process of construction, simplification, and sanitization occurs in the formation and use of thoughts and memories.⁷⁰ In this way, the illusion of an intentional self offers a simplified and sanitized sense of personal continuity that is nonetheless flexible enough to absorb the more complex and less sanitized aspects of life as they become relevant.⁷¹

As we saw in the previous chapter, Fonagy described the incoherence that confronts us when we try to understand others through attention to their every behavioural cue. An individual's concern with 'fleeting but genuine expressions of momentary emotional states is ultimately a hinderance rather than a help in their attempt to navigate complex social relationships', because other people are 'a theatre for too many states of mind for a coherent understanding of a relationship to be reliably achieved by these means.'⁷² The simplification and sanitization of the other's desires and experiences into 'thoughts' and 'feelings' implicated in motivations is therefore adaptive in allowing us to conceive of mental states in the other that can be used to account for their behaviour and infer their experience. To take an example: in loving, good-enough relationships, our attachment figures sometimes wish to kill us. This is normal. However, these wishes are usually fleeting, inchoate, and should receive acknowledgements and expression in no more than the tone or manner of safe and unthreatening responses.⁷³ For instance, responsiveness may be coloured by exasperation.

When things are well, such momentary expressions are also contextualized by what happens next, as well as the wider context of the relationship. It is notable that, when coding disorganized attachment, the guidance from Mary Main is to discount brief displays of fearful, conflicted, or confused behaviour 'if the baby goes to the parent as though for comfort after a bit of disorganisation.'⁷⁴ A degree of simplification and sanitization is fine when past experience has led us to basically conclude that our needs will be acknowledged and, where possible, met. If the vile passing thoughts of our attachment figures do not hinder their reflective functioning, Fonagy and colleagues assume that they should not hinder our capacity to make use of attachment figures as a secure base and safe haven as needed as children and

⁶⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'Emotions, which are the individual's bodily reaction, as action programs, to specific stimuli. Feelings, which are the conscious experience of the body state during emotional activation' (p. 307).

⁷⁰ See Fonagy, P. (1991). 'Thinking about Thinking: Some Clinical and Theoretical Considerations in the Treatment of a Borderline Patient'. *The International Journal of Psychoanalysis*, 72: 639–656; Fonagy, I. (1999). 'The Process of Remembering: Recovery and Discovery'. *The International Journal of Psychoanalysis*, 80(5): 961–978; Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, Chapter 6.

⁷¹ Luyten, P., Campbell, C., and Fonagy, P. (2020). 'Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach'. *Journal of Personality*, 88(1): 88–105; 'The self, and particularly the sense of self-coherence and self-continuity over time, is an illusion (Bargh, 2011, 2014) that is the product of the capacity for reflective functioning or mentalizing' (p. 91).

⁷² Fonagy, P. (2011). 'Discussion of Juan Pablo Jimenez's Paper, "A Fundamental Dilemma of Psychoanalytic Technique. Reflections on the Analysis of a Perverse Paranoid Patient"', in J. P. Jimenez and R. Muguillansky (eds), *Clinical and Theoretical Aspects of Perversion*, London: Karnac Books, pp. 63–76, p. 71.

⁷³ Winnicott, D. W. (1949). 'Hate in the counter-transference'. *The International Journal of Psychoanalysis*, 30: 69–74. See the discussion in Higgitt, A. and Fonagy, P. (1992). 'Psychotherapy in Borderline and Narcissistic Personality Disorder'. *The British Journal of Psychiatry*, 161(1): 23–43, p. 30.

⁷⁴ Main, M. (undated). *Disorganised/Disoriented Classification Scheme: Major Considerations*, unpublished manuscript. Received from Elizabeth Carlson, and cited with her permission.

as adults. The ability to use social relationships as a secure base and safe haven for help in processing difficult or confusing thoughts and feelings was theorized by Fonagy and colleagues to facilitate our capacity to conceive of and think about our thoughts and feelings in a coherent and consistent way.⁷⁵

The implication of the argument presented by Fonagy and colleagues is that mentalization of one's own and others' mental states genuinely, even if only to an extent, *creates* the coherence and consistency of thoughts, feelings, and/or intentions that it appears to merely represent. What partial personal coherence and consistency an individual has as an experiencing embodied subject, as a knower of themselves, as a personality, as an agent, as a social identity, or someone able to distinguish between personally authentic and inauthentic action—such attributions to the 'self' are theorized by Fonagy and colleagues as consequences, not primary causes, of mentalizing. As Bateman and Fonagy have argued: 'our sense of personal continuity is dependent on envisioning the thoughts and feelings we had in the past and how these relate to our current experiences ... mentalising, the representation of our mental states, is the spine of our sense of self and identity'.⁷⁶ The image of mentalizing as a spine for the self seems a rather essentializing metaphor, and reflects a tendency in the work of Fonagy and colleagues to reify the concept. Nonetheless, the intention of Bateman and Fonagy appears to be to highlight that mentalization and non-mentalizing are very important for the kind of self-representation a person develops. As such, mentalization in turn is not treated as free-standing, but as contextually embedded. Fonagy and colleagues also acknowledge that the characteristics of our self-representations do in turn 'shape the way people mentalize themselves' by forming 'a kind of heuristic for the individual making sense of his or her actions'.⁷⁷

For instance, Fonagy has stated that it is a particular point of alignment between mentalization-based therapy and cognitive analytic therapy to see that 'deficits in such reflective self representations are causal in stunting the development of a sense of oneself as an effective agent'.⁷⁸ However, he disagreed with cognitive analytic therapy, which he regarded as holding that once a representation of self is assembled, it is then passive.⁷⁹ Fonagy and colleagues have argued that the self-representation contributes actively to the process of imagining mental states in oneself and others, especially once the strengthened capacity for abstract representation of mental states becomes potentially available in the course of adolescence.⁸⁰ For instance, an inchoate wish or experience may be surfaced and elaborated

⁷⁵ However, major separations or the fear or threat of separations from attachment figures may, nonetheless, threaten to disrupt these benefits. Fonagy, P. and Moran, G. S. (1990). 'Severe Developmental Psychopathology and Brittle Diabetes: The Motivation for Self-Injurious Behaviour'. *Bulletin of the Anna Freud Centre*, 13: 231–248: 'internal danger ... is invariably associated with the psychic distance of the object, felt by the patient to be either invading the self or abandoning it. In either case what is feared and anticipated is the destruction of the self. This is because it is the coherence and continued stability of the mental representation of the self that comes under threat either from unwelcome affect-laden transactions with a dangerous object or enforced or violent separation from a desired, protective one' (p. 233).

⁷⁶ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 5.

⁷⁷ Fonagy, P., Campbell, C., and Allison, E. (2019). 'Therapeutic Models Mentalising', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 171.

⁷⁸ Fonagy, P. (1995). '2: Peter Fonagy'. *British Journal of Psychotherapy*, 11(4): 575–584, p. 576.

⁷⁹ The position of Fonagy and colleagues here runs contrary to that of Meissner, who held that 'if the self is merely and exclusively representational, it cannot serve as a source of action or agency ... it fails the demands of the role of the self in subjectivity and personal agency' Meissner, W. W. (1986). 'Can Psychoanalysis Find its Self?' *Journal of the American Psychoanalytic Association*, 34(2): 379–400, p. 382.

⁸⁰ See Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications: 'Perhaps Sandler's most important contribution was his distinction between the experiential and the non-experiential realms. Whereas the former referred to Sandler and Joffe's

into a thought or feeling as a mental state due to its relevance to the concerns of the self-representation—perhaps its pertinence to personally relevant tasks, projects, or worries.

This perspective adds a deeply uncanny quality to Fonagy's seemingly benign claim that caregiving that acknowledges the child as an agent supports the child's representation of a 'constitutional self'. This constitutional self is not, as it might appear at first sight, the 'authentic' and 'true' self of Rogers (or, sometimes, Winnicott).⁸¹ For Fonagy, an authentic and true focal awareness of our own chaotic and unnerving desires would be both authentically disturbing and truly obstructive. Instead, the appearance of a 'constitutional self' out of the complexity of experience is itself an illusion. Where formed through mentalization, it is an illusion that is resilient to breakdown because its coherence and consistency are based around the capacity to conceive of thoughts and feelings, and the capacity to reconsider them when more complexity needs to be considered. The main forms of non-mentalizing—pretend mode, psychic equivalence, and teleological mode—also offer the illusory coherence and consistency needed to inform action. This is part of what makes them adaptive under particular circumstances (see Chapter 7). However, the kind of coherence and consistency they offer blocks access to the reconsideration of thoughts or feelings. In fact, Fonagy and Target proposed, they can make the generative doubt that permits reconsideration feel as a kind of 'mutilation', because a sense of personal coherence and consistency is sustained only by excluding such doubt.⁸²

representational model, the latter entailed mechanisms, structures and apparatuses. The non-experiential is inherently non-conscious, although it is not repressed or dynamically inhibited. The distinction between a fantasy (conscious or unconscious) and the organized function underpinning it (fantasizing) remains an evocative example. The model makes clear that experience is not the agent of change; rather, change is brought about by structures in the non-experiential realm, which cause corresponding changes in the experiential. Thus self-representation cannot be an agent, but it is an entity that will determine how mechanisms of the mind behave ... Sandler (1990) clarified his view of internal objects as "structures" within the non-experiential realm, albeit constructed out of subjective experience, conscious or unconscious. Once created, such non-experiential structures can modify subjective experience, including the child's experience of actual objects (people)' (p. 105). On adolescence as both a sensitive period and a challenging period for the solidification of a self-representation through the capacity to imagine mental states in oneself and others, see Sharp, C. and Rossouw, T. (2019). 'Borderline Personality Pathology in Adolescence', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalisation in Mental Health Practice*, Washington, DC: American Psychiatric Association, pp. 281–300. Similar ideas were presented already across Anna Freud's work, and may have provided a background influence for Fonagy and colleagues on this point.

⁸¹ The case of Winnicott, always a little tricky to pin down as a theorist, is more complicated than Rogers. When he introduced the concept, Winnicott urged 'The True Self appears as soon as there is any mental organization of the individual at all, and it means little more than the summation of sensori-motor aliveness': Winnicott, D. W. ([1960] 1965). *Ego Distortion in Terms of True and False Self, the Maturation Process and the Facilitating Environment*, New York: International Universities Press, pp. 140–152, p. 149. This would not be illusion in Fonagy's terms. However, in Winnicott's later writings, the term is explicitly aligned with the idea of authenticity in Romanticism, and takes a more essentialist cast, which draws the concept back towards Rogers and the target of Fonagy's critique. Winnicott, D. W. (1986). 'The Concept of the False Self', in *Home is Where We Start From*, New York: W. W. Norton, pp. 65–70. The potential for Winnicott's concept of 'true self' to slide towards essentialism, and even moralizing essentialism, has been criticized by some commentators. However, Winnicott has also been defended by others on this count. See e.g. Papadima, M. (2006). 'Dissociation, the True Self and the Notion of the Frozen Baby'. *Psychodynamic Practice*, 12(4): 385–402; Ruti, M. (2010). 'Winnicott with Lacan: Living Creatively in a Postmodern World'. *American Imago*, 67(3): 353–374.

⁸² Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration'. *Bulletin of the Anna Freud Centre*, 17: 33–55: 'The analytic goal must be to engender within the patient a true understanding. Understanding here implies detachment, i.e. the permanent loss of the omnipotent self (Gaddini, 1987). To admit to analytic understanding is equivalent to recognizing extreme fragility and vulnerability of a self which appears mutilated in comparison to the self that is proximal to the understanding object, however distorted that understanding may be' (p. 53). The term 'mutilation', and in fact much of the wording of this statement, is drawn from Gaddini, E. ([1980] 1992). 'Notes on the Body-Mind Question', in Adam Limentani (ed.), *A Psychoanalytic Theory of Infantile Experience: Conceptual and Clinical Reflections*, London: Routledge, pp. 105–124, p. 115.

In a chapter from 2012, Fonagy, Bateman, and Luyten interrogated the role of 'self' in forms of mental disorder, distinguishing between three kinds of disorders based on the failures of the illusion of personal consistency and coherence they represent.⁸³ Disorders of the 'integrity of self', chief among them BPD, are caused by sustained use of non-mentalizing modes of processing, because these hinder the creation of durable, flexible, and adaptive illusions of individual intentionality as implicated in thoughts and feelings. Fonagy and colleagues described 'disorders of self-recognition', whereby an individual's embodied experience has become decoupled from the appearance of personal relevance, a stance characteristic of pretend mode. Capgras' syndrome (a condition in which familiar people are delusionally experienced as strangers) may be an extreme form, but Fonagy and colleagues proposed that self-harming behaviours also reflect a disorder of self-recognition. Third, 'disorders of self-concept and self-image' are prompted when non-mentalizing specifically distorts the simplification and sanification of embodied experience, with the result that incoherent or disturbing ideas or feelings, which would otherwise be fleeting or just colour perception, can become pinned as stable aspects of the individual's perception of their own state and worth. Skårderud and Fonagy described the case of patients with eating disorders.⁸⁴ The culture of Western societies symbolizes willpower and self-control in terms of control over weight. This predisposes teleological thinking in which the normal, passing feelings of distrust, bafflement, and frustration individuals experience may be noticed, ruminated on—and then mistaken as caused by the individual's body fat. Mental states are read off perceptions of physical qualities rather than physical cues being used to inform mental states.

That non-mentalizing can undermine the illusion of self-representation, and its coherent and consistent functioning, has significant implications for the goals of clinical work. A first point is that Fonagy has disagreed with those for whom 'the main developmental achievement for any individual pertains to the achievement of a cohesive self'.⁸⁵ This is a position he attributes especially to Anthony Ryle and cognitive analytic therapy, though he considers that it characterizes various forms of short-term counselling and psychotherapy. The problem with this approach, Fonagy argued, is that coherence and consistency are an illusion. To direct the client to seek a cohesive self risks failing to help them develop strategies for dealing with the inevitable, all-too-human aspects of the mind that exceed this coherence, and that at times disrupt or simply run contrary to our self-representation. Among these, for example, are experiences of desire and frustration, as we shall discuss shortly. Fonagy worried that a focus on achievement of a cohesive self 'may re-enforce the splitting, rather than aid the individual to develop a capacity to tolerate ambivalence, intense affect and, in the long term, reduce impulsivity and the marked oscillation of self esteem'.⁸⁶

⁸³ Fonagy, P., Bateman, A. W., and Luyten, P. (2012). 'Introduction and Overview', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 3–42, p. 7.

⁸⁴ Skårderud, F. and Fonagy, P. (2012). 'Eating Disorders', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 347–384: 'We will first describe the psychopathology of eating disorders by using the language of the mentalizing model. We understand these disorders as manifestations of an underlying self-disorder. This underlying disorder should be the central focus of psychotherapy' (p. 348). See also Sacchetti, S., Robinson, P., Bogaardt, A., Clare, A., Ouellet-Courtois, C., Luyten, P., ... and Fonagy, P. (2019). 'Reduced Mentalizing in Patients with Bulimia Nervosa and Features of Borderline Personality Disorder: A Case-Control Study'. *BMC Psychiatry*, 19(1): 134.

⁸⁵ Fonagy, P. (1995). '2: Peter Fonagy'. *British Journal of Psychotherapy*, 11(4): 575–584, p. 580.

⁸⁶ Ibid. Ryle's riposte to Fonagy was to claim that the sense of self is constituted by a person's interpersonal and intrapersonal social experiences, not merely a representation of that experience. On this basis, Ryle and Kerr regarded the integration of a person's interpersonal and intrapersonal social experiences as realistic and worthwhile, even if it is a task that is never complete. Ryle, A. and Kerr, I. B. (2003). *Introducing Cognitive Analytic Therapy: Principles and Practice*, New York: Wiley, p. 37.

Nonetheless, Bateman and Fonagy listed several ways that supporting a patient's capacity to mentalize and to hold off from non-mentalizing processing should contribute to greater coherence and consistency of the self-representation.⁸⁷ First, improved mentalizing and less use of non-mentalizing modes can be anticipated to heighten the viability, clarity, and acuity of conceptions of the thoughts and feelings of oneself and others, serving to inform and stabilize the self-representation.

Second, Bateman and Fonagy stated that this can help the 'formation of a coherent sense of self'.⁸⁸ Here they slip back into a cluttered use of the term 'self' that elsewhere they have criticized in others. From comparison of the Bateman and Fonagy paper with other work using similar language,⁸⁹ the meaning would appear to be either i) that mentalization can be anticipated to contribute to the coherence and consistency of an individual's experience as an embodied subject; or ii) that mentalization can be anticipated to contribute to a person's sense of coherent personal agency. Or both meanings may have been intended. On the one hand, mentalization can be expected to help an individual sift their lived experience for making coherent thoughts and feelings, and also for making them available for pruning and revision, contributing further to coherence and consistency. On the other hand, even if the image of ourselves as agents is, in Fonagy's view discussed above, an illusion, he has described it as a helpful illusion that is partly made real by its enactment. When a person acts with a sense of viable personal agency, this helps organize experience by helping to support and refine relevant thoughts and feelings.

Third in Bateman and Fonagy's list, they proposed that mentalization and avoidance of non-mentalizing can also be expected to contribute to the 'capacity to form secure relationships'.⁹⁰ As we have seen, the ability to use social relationships as a secure base and safe haven for processing difficult or confusing thoughts and feelings was theorized by Fonagy and colleagues as helping to conceive thoughts and feelings and make them available for reflective function.

Finally, Bateman and Fonagy observed that emotional turmoil is likely to disrupt the self-representation and its contribution to the construction, identification, and use of relevant mental states. When a chronic and reciprocal disturbance occurs between emotional turmoil, mentalizing, and social relationships, the result is what has been called a personality disorder, or what Bateman and Fonagy described as a 'destabilisation' of 'the self-structure',⁹¹ though the exact meaning of the concept of 'self-structure', and its place within the metapsychology of mentalization theory, remains uncertain. The term originates in Rogers' humanistic psychotherapy as a synonym for personality, but it is not clear whether or not Bateman and Fonagy intended it in this sense.⁹²

⁸⁷ Bateman, A. W. and Fonagy, P. (2003). 'The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder', *Bulletin of the Menninger Clinic*, 67: 187–211, p. 195.

⁸⁸ Ibid.

⁸⁹ Fonagy, P. and Target, M. (1997). 'Attachment and Reflective Function: Their Role in Self-Organization', *Development and Psychopathology*, 9(4): 679–700.

⁹⁰ Bateman, A. W. and Fonagy, P. (2003). 'The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder', *Bulletin of the Menninger Clinic*, 67: 187–211, p. 195.

⁹¹ Ibid.

⁹² Rogers, C. R. (1959). *Client-Centered Therapy: Its Current Practice, Implications, and Theory*, Boston: Houghton Mifflin. The term's first sustained use within psychoanalysis was in conceptualizing borderline and narcissistic personality disorders—for instance, in the work of Irene Fast, as well as Kernberg and then Kohut. The first explicit definition of the term appears to be in Frances, A., Sacks, M., and Aronoff, M. S. (1977). 'Depersonalization: A Self-Relations Perspective', *The International Journal of Psychoanalysis*, 58: 325–331: 'We use the term "self-structure" to describe the coherent organization of those previously registered self representations which provide an individual with his experiential sense of psychological intactness, i.e. his sense of self' (p. 325).

Externalizing the alien self

As we have seen, Fonagy and colleagues characterize the coherence of self-representation as an illusion, but an adaptive one, because it contributes variously to the coherence and consistency of embodied experience, self-knowledge, self-structure, and personality, personal agency, social identity, and feelings about personal authenticity. However, for Fonagy and colleagues, this coherence and consistency are only ever partial, even under the best of circumstances. They remain vulnerable to disruption by the states of fragmentation and meaninglessness of the primary unconscious. And they remain vulnerable to the pull of the contrary and disturbing intentions of the psychoanalytic unconscious. Caregiver reflective function will support greater emotional containment and acknowledgement of a child's intentionality and/or mental states, facilitating the coherence and consistency of the child's emerging self-representation and their capacity to flexibly adapt this representation through mentalization. The result is 'an authentic, organic self-image built around internalised representations of self-states.'⁹³ However, total containment and acknowledgement are impossible; in fact, Fonagy has speculated that the fantasy of an omniscient mentalizing divine being may stem from this disappointment.⁹⁴

All children develop a psychoanalytic unconscious, which for Fonagy and colleagues comprises intentions (in the sense of wishes and plans) with reference to the self, others, or the world that have been refused acknowledgement, and excluded from the self-representation.⁹⁵ The result is the experience of an 'alien self', the experience of intentions that do not agree with the self-representation, and so do not feel like they belong to us.⁹⁶ The idea of the 'alien self' was not original to Fonagy and colleagues, but stemmed from second-generation Kleinian refinements of the concept of projective identification. In the most general terms, projective identification means to imagine some part of oneself as able to be split off, and to expel it outside and away—for instance, into the environment or another person.⁹⁷ The term 'alien self' was used by second-generation Kleinians to describe the individual's experience of such disowned aspects of the self.⁹⁸ For Kleinians, some of these

⁹³ Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873, p. 864.

⁹⁴ Fonagy, P. (2009). 'Commentary on "Forgiveness"', in S. Akhtar (ed.), *Good Feelings: Psychoanalytic Reflections on Positive Emotions and Attitudes*, London: Karnac Books/International Psychoanalytic Association, pp. 411–452. 'In imagining understanding someone else fully, or being fully and accurately understood by them, we are forced into the world of imagination, the fantasy of an omniscient mentalising being' (p. 423).

⁹⁵ The elaboration of unacknowledged experiences into intentions in the psychoanalytic unconscious is described, though only briefly and speculatively, in Fonagy, P. and Allison, E. (2016). 'Psychic Reality and the Nature of Consciousness'. *The International Journal of Psychoanalysis*, 97(1): 5–24. 'Although it is "unmetabolized", the psychoanalytic unconscious acquires a partial quality of intentionality or "aboutness" through the process of projective identification. When aspects of sexuality or destructiveness are communicated by the infant to the mother, she does not mirror but recognizes the experiences, transmitting them back infused with intentionality from her unconscious associations to the feelings' (p. 13).

⁹⁶ Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 439.

⁹⁷ Goretti, G. R. (2007). 'Projective Identification: A Theoretical Investigation of the Concept Starting from "Notes on some Schizoid Mechanisms"'. *The International Journal of Psychoanalysis*, 88: 387–406.

⁹⁸ On their first usage, Fonagy and Target attribute the idea to Britton's 1998 book on the imagination. However, Britton in turn was drawing on Grotstein. See Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873; Britton, R. (1998). *Belief and Imagination*, London: Routledge; Grotstein, J. S. (1977). 'The Psychoanalytic Concept of Schizophrenia: II. Reconciliation'. *The International Journal of Psychoanalysis*, 58: 427–452; Grotstein, J. S. (1980). 'A Proposed Revision of the Psychoanalytic Concept of Primitive Mental States: Part I. Introduction to a Newer Psychoanalytic Metapsychology'. *Contemporary Psychoanalysis*, 16(4): 479–546. The 'alien self' was

aspects are wishes and plans disowned as too positive for the self-representation, forming an 'ego ideal'; Fonagy and colleagues did not discuss these.⁹⁹ Their account of the 'alien self' focused on those aspects that are disowned as unacceptable to the self-representation. They proposed a developmental account of the emergence of this alien self:

Even the most sensitive caregiver is insensitive to the child's state of mind more than 50% of the time. Thus, we all have alien parts to our self-structure. The coherence of the self, as many have noted, is somewhat illusory. This illusion is normally maintained by the continuous narrative commentary on behaviour that mentalisation provides, which fills in the gaps and weaves our experiences together so that they make sense. In the absence of a robust mentalising capacity, with disorganised patterns of attachment, the fragmentation of self-structure is clearly revealed.¹⁰⁰

For Fonagy and Target, then, mentalization is partly a confabulation: the production of a somewhat fictionalized story, woven into the experience of an individual 'self'. In this, their position aligns with that of Winnicott, for whom the binding of the experience of self to an individual bounded body was dependent upon an 'imaginative elaboration'. However, the reference by Fonagy and Target to 'narrative commentary' hints, further, that this is not just an elaboration based on 'body functioning of all kinds and the accumulations of memories', but also of available cultural narratives for integrating and reconciling gaps and experiences of being alive within a particular context.¹⁰¹ Though not explicit in Winnicott, this is well in line with his account of the importance of environmental factors and social communication in the formation of a sense of self.¹⁰² This role of cultural narratives in the imaginative elaboration of the self-representation and the alien self was not elaborated at the time by Fonagy and Target, conveying the impression that there would be little cross-cultural variation in the formation of the self as an individual. The acknowledgement of the self as a confabulation was not accompanied by recognition of contingency in the construction of personhood. This had the side-effect of contributing to a naturalization of individualism within mentalization theory (see Chapter 2) precisely at the point that the individual self was theorized as a narrative construction and, to an extent, confabulation.¹⁰³ As we shall see in Chapter 9, Fonagy

predated in psychoanalytic theory by the idea of 'ego-alien feelings', which was a definite precursor, but lacked the implication of a sense of agency or intentionality that was within a person without belonging to them.

⁹⁹ The reason for the neglect of positive introjects by Fonagy and colleagues is not clear. It could be that they assumed that positive introjects simply form part of the self-representation. Or it could be that they regarded positive introjects as of less clinical and developmental relevance. As a point of comparison, highlighting the importance of positive projective identification for caregiving, see Likierman, M. (1988). 'Maternal Love and Positive Projective Identification'. *Journal of Child Psychotherapy*, 14(2): 29–46. It may also be that positive introjects can contribute, precisely by feeling somewhat other than the self, to the successful calibration of epistemic vigilance (see Chapter 7)—for instance, through establishing an effective ego ideal.

¹⁰⁰ Fonagy, P. and Target, M. (2003). 'Evolution of the Interpersonal Interpretive Function: Clues for Effective Preventive Intervention in Early Childhood', in S. W. Coates, J. L. Rosenthal, and D. S. Schechter (eds), *September 11: Trauma And Human Bonds*, Hillsdale, NJ: Analytic Press, pp. 99–113, p. 106. See also Luyten, P., Campbell, C., and Fonagy, P. (2020). 'Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach'. *Journal of Personality*, 88(1): 88–105.

¹⁰¹ Winnicott, D. W. (1990). *Human Nature*, London: Routledge, p. 23.

¹⁰² See e.g. Winnicott, D. W. (1968). 'Communication between Infant and Mother, and Mother and Infant, Compared and Contrasted', in Walter G. Joffe (ed.), *What is Psychoanalysis?*, London: The Institute of Psychoanalysis/Ballière, Tindall & Cassell; Winnicott, D. W. (1971). *Playing and Reality*, London: Tavistock. Cf. Hutto, D. D. (2008). 'The Narrative Practice Hypothesis: Clarifications and Implications'. *Philosophical Explorations*, 11(3): 175–192.

¹⁰³ Cf. Ganeri, J. (2012). *The Self: Naturalism, Consciousness, and the First-Person Stance*, Oxford: Oxford University Press; Ganeri, J. (2017). *Attention, Not Self*, Oxford: Oxford University Press.

and colleagues have recently sought to amend and advance their position with greater attention beyond the individual and to questions of cultural variation.

For Fonagy and Target, the particular importance of the imaginative integration of the self-representation lay in its contribution to the experience of coherence and consistency through supporting mentalization and the capacity to form secure attachments. By contrast, a more accurate characterization of our experiences, intentions, self-knowledge, etc. is of 'fragmentation', a state 'revealed' when mentalization fails to generate confabulation and its ensuing coherence (see Chapter 3). The intentions of the psychoanalytic unconscious, let alone the meaninglessness of the primary unconscious, are disturbing and disruptive. As a consequence, any truly *accurate and unmediated* reflection of these in our behaviour would indeed be fearful, conflicted, and confused to an extent. The developmental importance of caregiver reflective functioning, in Fonagy and Target's account, is that it supports the development of a mental apparatus for simplifying and sanitizing experience to permit the formation of coherent thoughts and feelings, capable of being subject to reconsideration. This apparatus can even offer protection in the context of later adversity and trauma: mentalization can fill in the gaps and weave our experiences together enough to make sense.

Yet 'in the absence of a robust mentalizing capacity, in the wake of trauma, alien fragments in the self-structure are likely to be clearly revealed in all of us'.¹⁰⁴ In the 2000s, Fonagy and Target situated poor caregiver reflective functioning as an important developmental basis for a strong alien self. The child is unable to find the basis of their coherent self-representation and intentionality in their caregiver's responses. This leads to the 'internalisation of representations of the parent's state, rather than of a usable version of the child's own experience'.¹⁰⁵ These states may not be benign ones and may be further distorted by the child's own unassuaged concerns if they are unable to use their caregiver as a secure base and safe haven for difficult thoughts and feelings. In this way, experiences of non-acknowledgement by attachment figures may form 'the germ of a potentially persecutory object which is lodged in the self, but is alien and unassimilable'.¹⁰⁶

In recent years, Fonagy and colleagues have argued for an important role for later experiences, after childhood, in the intensification of the experience of an alien self. In the absence of mentalization, traumatic experiences—for instance, political violence and torture—reduce an individual's capacity to inhibit conflicting appraisals and to buffer the effects of emotional reactions on information processing.¹⁰⁷ This hinders the individual's capacity to sift coherent and consistent thoughts and feelings, and exclude intentions from the psychoanalytic unconscious incompatible with the self-representation.¹⁰⁸ Fonagy and colleagues

¹⁰⁴ Fonagy, P. and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology'. *Journal of Personality Disorders*, 20: 544–576, p. 567.

¹⁰⁵ Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873, p. 864. See also Fonagy, P., Gergely, G., Jurist, E. L., and Target, M. (2002). *Affect Regulation, Mentalization, and the Development of the Self*, London: Karnac Books.

¹⁰⁶ Fonagy, P. and Target, M. (1999). 'Towards Understanding Violence: The Use of the Body and the Role of the Father', in R. J. Perelberg (ed.), *Psychoanalytic Understanding of Violence and Suicide*, London: Routledge, pp. 53–72, p. 62.

¹⁰⁷ Lorenzini, N., Campbell, C. and Fonagy, P. (2019). 'Mentalisation and its Role in Processing Trauma', in Bernd Huppertz (ed.), *Approaches to Psychic Trauma: Theory and Practice*, Lanham, MD: Rowman & Littlefield, pp. 403–422.

¹⁰⁸ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 1. Borderline Personality Disorder as a Limitation of Resilience'. *Borderline Personality Disorder and Emotion Dysregulation*, 4: 11.

have also speculated that wider dehumanizing culture and institutionalized values may intensify the experience of the alien self (see Chapter 9).

The experience of an alien self 'interferes with the relationship between thought and identity: ideas or feelings are experienced that do not seem to belong to the self. The alien self destroys the sense of coherence of self.'¹⁰⁹ The feeling of coherence within perceptual experience, self-knowledge, and agency may be achieved through alternate means, in the use of pretend mode, psychic equivalence, or teleological mode. However, Fonagy and colleagues also consider another strategy that restores feelings of coherence, which stands as the closest candidate in their writings to a distinct, 'fourth' major form of non-mentalizing.¹¹⁰ This is termed 'externalization of the alien self'. In fact, Luyten, Malcorps, Fonagy, and Ensink list externalization of the alien self alongside pretend mode, psychic equivalence, and teleological mode in their table of non-mentalizing in the latest edition of the *Handbook of Mentalisation in Mental Health Practice*.¹¹¹ This could be regarded as a hint that its elevation to the official status of a mode of non-mentalizing may be under tacit consideration. However, declining reference to the concept in recent years would suggest otherwise, and instead point to the rather unstable and poorly integrated status of the concept for Fonagy and colleagues.

In the externalization of the alien self, 'the other's mind is being controlled or manipulated in order to restore a sense of self in the face of an assault that has inhibited the capacity to regard one's own mind as coherent. This does not represent a complete absence of mentalizing; rather, it is a form of highly disrupted mentalizing in which the other's mind is recognized and used to restore one's sense of one's own mind.'¹¹² Externalization of the alien self is a revised account of the classic but overlaid Kleinian concept of 'projective identification'. This may have slowed its acknowledgement as a distinct form of non-mentalizing, because there was already an apparently encompassing term, and one with a very complicated and messy history.¹¹³ This would have made attempts to excavate and redefine the concept hazardous for Fonagy and colleagues in the 1990s and early 2000s, not only because the work would have been fiddly and technical, but also because it could have appeared as a capitulation to a

¹⁰⁹ Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873, p. 864.

¹¹⁰ E.g. Bateman, A. and Fonagy, P. (2008). 'Comorbid Antisocial and Borderline Personality Disorders: Mentalization-Based Treatment'. *Journal of Clinical Psychology*, 64(2): 181–194: 'Stability is maintained by interpreting the world according to teleological understanding much of the time, but in many instances, we suggest that stabilization of mental processes arises from the rigidity of the externalization of the alien self' (p. 184).

¹¹¹ Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Mentalising and Trauma Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 79–102, p. 85, Table 5.1. Luyten and colleagues also hint at 're-enactment of trauma' as a potential additional form of non-mentalizing, though it might be supposed that this could be accounted for in terms of the other three—or four—modes of non-mentalizing. The relationship between re-enactment of trauma and the forms of non-mentalizing is not addressed in depth by the authors.

¹¹² Asen, E. and Fonagy, P. (2017). 'Mentalizing Family Violence Part 1: Conceptual Framework'. *Family Process*, 56(1): 6–21, p. 13.

¹¹³ Sandler, J. (ed.) (1988). *Projection, Identification, Projective Identification*, Karnac Books. Bateman, A. W., and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'What psychodynamic clinicians would recognise as "projective identification". This term has many meanings, and this has led us to talk about one aspect of this—the externalisation of the alien part of the self' (p. 20). The idea of 'externalization' was developed by Anna Freud and Joseph Sandler in responding to Klein's work, again attempting to distinguish specific meanings within the overburdened concept of projective identification. By 'externalization', Freud and Sandler intended the ascription of a troubling mental state of one's own to someone or a group in the outside world, a defence mechanism they regarded as capable of becoming pathological if implemented too rigidly and intensely, but also as quite ordinary and sometimes functional in human life. Sandler, J., Kennedy, H., and Tyson, R. L. (1975). 'Discussions on Transference: The Treatment Situation and Technique in Child Psychoanalysis'. *The Psychoanalytic Study of the Child*, 30(1): 409–441; Sandler, J. and Freud, A. (1981). 'Discussions in the Hampstead Index on "The Ego and the Mechanisms of Defence": IV. The Mechanisms of Defence, Part I'. *Bulletin of the Anna Freud Centre*, 4(3): 151–199.

Kleinian perspective at a point that Fonagy was being accused of being insufficiently Anna Freudian (see Chapter 1).¹¹⁴ The relatively insular London Kleinian group had taken up projective identification as their flagship concept in the 1990s, giving it a specific symbolic load in the immediate context in which Fonagy and colleagues were working.¹¹⁵

An outstanding problem with the concept of 'externalization of the alien self' is inherited from 'projective identification'. As Fonagy and colleagues observed in the 1990s, the concept of projective identification risks confusing two things. On the one hand, phenomenology: what it feels like to the individual. On the other hand, explanation: the psychological mechanism entailed.¹¹⁶ This is precisely the distinction elided by psychic equivalence, where experience is taken as explanation; a good part of mentalization-based therapies is helping patients think about their experience, and consider possible explanations.¹¹⁷ Fonagy and Target acknowledged that the concept of projective identification appeals precisely because it *feels* like it captures the way patients can get under the therapist's skin. But the reality of this feeling is distinct from a conceptualization of projective identification in terms of causal mechanisms.¹¹⁸ The same problem, however, has attended their uses of the concepts of 'mentalization' and 'disorganized attachment' (Chapter 3) as we have seen, and the concept of 'epistemic trust' as we will see in the next chapter (Chapter 7). It also attends their use of the terms 'self' and 'alien self'.

Sometimes Fonagy and colleagues characterize the alien self as an aspect of an individual's experience. Sometimes they treat it as a psychological structure. Sometimes it appears to be unconscious, outside of an individual's experience. Sometimes it is treated as an aspect of all

¹¹⁴ See e.g. Bach, S., Alvarez, A., Mayes, L., and Fonagy, P. (2000). 'Panel 3: Fantasy Life and the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 51–62. 'Peter Fonagy: I always had terrible trouble with the notion of projective identification. I was trained at the Anna Freud Centre and we just didn't talk about things like that' (pp. 60–61).

¹¹⁵ Hargreaves, E. (2004). *In Pursuit of Psychic Change: The Betty Joseph Workshop*, London: Routledge. See also Baert, P. (2012). 'Positioning Theory and Intellectual Interventions'. *Journal for the Theory of Social Behaviour*, 42(3): 304–324. See also Brown, L. J. (2012). 'The New Dictionary of Kleinian Thought. By Elizabeth Bott Spillius, Jane Milton, Penelope Garvey, Cyril Couve, and Deborah Steiner (Review)'. *Psychoanalytic Quarterly*, 81(3): 775–780: 'Some years ago, I attended a symposium on the Controversial Discussions, hosted by the Massachusetts Institute for Psychoanalysis, at which the featured speakers were Ron Britton (representing the Kleinian view), Peter Fonagy (the contemporary Freudian view), and David Tuckett (the Independent view). At one point, Fonagy spoke about the tendency toward insularity among Kleinians, and also acknowledged a measure of envy about that group's collective focus on systematically developing their concepts' (p. 775).

¹¹⁶ In the estimation of Fonagy and Hepworth, confusion of phenomenology and explanation characterized much of Klein's terminology. Technically, this would mean that much of Kleinian therapy would be conducted close to psychic equivalence, a state in which phenomenology is taken for explanation. Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications: 'Criticisms address the "fuzziness" of Klein's descriptions. The emphasis on "phantasy" as the building block of mental structure means that mental structuralization has been moved into the experiential realm (see Sandler & Joffe 1969), rather than being seen as inaccessible to awareness. This carries the advantage of closeness to clinical experience, and rids theory of much reified pseudo-scientific terminology. However, it bypasses essential questions and findings about the mechanisms underpinning mental functions.' (135, parentheses suppressed); 'The attainment of the "depressive position" illustrates some of the ambiguities of Kleinian terms. This change (whether or not seen as a stable development stage) clearly implies a qualitative shift in the perception of the object from part to whole. It is not clear, however, if it also implies a) consciousness of conflicted feelings about the same person (e.g. love and hate); b) the unconscious integration of various images with no necessary conscious correlate; c) the ability to recognize that the same person can generate conflicting feelings, but that these do not necessarily "belong" to that person' (p. 136).

¹¹⁷ E.g. Lemma, A., Target, M., and Fonagy, P. (2011). *Brief Dynamic Interpersonal Therapy: A Clinician's Guide*, Oxford: Oxford University Press: 'A DIT formulation has several components: 1) It describes the problem as seen by the patient. 2) It contextualizes the problem in a developmental framework ... 3) It pulls together this information into an account that meaningfully links the patient's difficulties with a psychological, dynamic process' (pp. 111–112).

¹¹⁸ Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 135

psychological processes to a greater or lesser degree. All of these are possibilities of course. A confusion between phenomenology and explanation has meant that Fonagy and colleagues have never really sorted these matters out or articulated their interrelation. Fonagy and colleagues tend to lack clarity on whether the feeling of externalizing the alien self is a fantasy and/or a kind of psychic equivalence, or whether something is actually externalized.¹¹⁹ If the former, the actual mechanisms that allow the other's mind to be controlled or manipulated would need to be articulated, as would the exact overlap or causal relationship between externalization of the alien self and psychic equivalence.¹²⁰ Kleinians such as Bion had identified this as a problem already for their concept of projective identification,¹²¹ and this issue appears to have been inherited by Fonagy and colleagues. Appeal to the alien self appears particularly in their more clinical publications. It rarely figures in their papers more concerned with underpinning mechanisms. Again, this may also have made it harder for externalization of the alien self to become identified as a fourth major mode of non-mentalizing, or for clarifying justification to be offered as to why it does not receive this status.

Like in pretend mode, psychic equivalence, and teleological mode, in externalization of the alien self—conceived of as a psychological mechanism—elements of mentalization are hijacked. *Present and past perceptual experience, observable social behaviour, and motivations and intentions of the other are conceived, accounted for and controlled in order to avoid conceiving of thoughts and feelings in oneself.* Bateman and Fonagy have described how 'if the alien self is an experience of vulnerability, the person creates this experience in his communication partner by generating chronic uncertainty; if it is aggression, he simply has to irritate him; if it is depression or lack of interest and hopelessness, then he might force him to experience the potential of helping, only to dash his hopes again and again'.¹²² Fonagy gave the autobiographical example of his panicked phone calls home to his parents as an adolescent living alone in London, where he would 'talk about his situation in catastrophic terms until his parents were palpably panicked, and then he would end the conversation feeling relieved'.¹²³ (see Chapter 1).

Illustrating the externalization of the alien self, Fonagy and Target described the case of Callum, who displayed aggressive and controlling behaviour when his mother was present, and intense separation anxiety when she was absent. Fonagy and Target acknowledged that an aspect of these symptoms expressed attachment anxiety about the availability of the caregiver, based on past experiences. Yet they argued that 'Callum's need for closeness to his mother is not simply motivated by anxiety but also by the need for a vehicle to contain the

¹¹⁹ Concerns about the lack of clarity in use of the concept of 'externalization' in this regard have been offered by Auerbach, J. S. (2005). Fonagy, P., Gergely, G., Jurist, E. L., and Targe, M. 'Affect Regulation, Mentalization, and the Development of the Self (Book Review)'. Accessed at: <https://www.apadivisions.org/division-39/publications/reviews/affect>.

¹²⁰ The possibility that the alien self is a kind of psychic equivalence is tantalizingly suggested in the first edition of Bateman and Fonagy's *Psychotherapy For Borderline Personality Disorder*. Though the link is not discussed in the text, in Figure 3.6, Bateman and Fonagy suggest that externalization of the alien self has a special causal relationship with psychic equivalence not possessed by teleological mode and pretend mode. However, the nature of this link is not elaborated. And the diagram disappears from the subsequent edition of the book. If externalization of the alien self is indeed partly a species of psychic equivalence, then this would suggest another reason for its exclusion as a fourth form of non-mentalizing.

¹²¹ Bion, W. ([1977] 2019). *Bion in New York and Sao Paulo, and Three Tavistock Seminars*, London: Karnac Books.

¹²² Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 21.

¹²³ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 21.

alien part of himself'.¹²⁴ In controlling his mother's behaviour, he also achieved an indirect form of self-control, permitting greater feelings of coherence.

Fonagy and colleagues have described the alien self as formed by our intentions that are not given acknowledgement, initially by caregivers and then later in other social interactions. Sexuality and aggression are therefore powerful components of the alien self of most humans. Fonagy and colleagues have argued that these are the forms of intentionality that receive the least acknowledgement from caregivers, and that have the most circumscribed place in social interaction, 'thus always remaining somewhat alien, and so prone to acting out',¹²⁵ though they acknowledge that there will be cultural variability in the degree and form of acknowledgement of these forms of intentionality. They suggest, for instance, that new social media may offer representations of sexuality and aggression that offer modes of part-acknowledgement and part-construction to sexual and aggressive intentionality (see Chapter 9). As a result, these aspects of the alien self may impinge more into the self-representation and its activities.¹²⁶ Nonetheless, Fonagy and colleagues have assumed that some externalization of the sexual and aggressive selves is a basic facet of human experience.¹²⁷ It is the externalization of the sexual and aggressive selves that gave earlier psychoanalytic theory, including Anna Freud, the impression that sexuality and aggression could be conceptualized as drives, characterized by intentions that in key regards exceed or transgress an individual's image of themselves.¹²⁸ The externalization of alien self in the cases of sexuality and aggression will be discussed in turn.

Sexuality and aggression

In Freud's classical stance, a child's sexual wishes are inherently unrealizable. The wishes are therefore excluded from consciousness but remain live in the unconscious. This creates a foundational conflict, implicated in various ways in mental health symptoms. For instance, in phobias, a forbidden sexual wish is regarded as linked to a particular object, which is then associated with fear and aversion. From the 1990s, Fonagy and Target expressed their agreement with Freud that sexual intentions are, at base, incongruent with self-representations.

¹²⁴ Fonagy, P. and Target, M. (2003). 'Evolution of the Interpersonal Interpretive Function: Clues for Effective Preventive Intervention in Early Childhood', in S. W. Coates, J. L. Rosenthal, and D. S. Schechter (eds), *September 11: Trauma And Human Bonds*, Hillsdale, NJ: Analytic Press, pp. 99–113, p. 110.

¹²⁵ Fonagy, P. and Campbell, C. (2015). 'Bad Blood Revisited: Attachment and Psychoanalysis, 2015'. *British Journal of Psychotherapy*, 31(2): 229–250, p. 241.

¹²⁶ Fonagy, P. and Allison, E. (2015). 'A Scientific Theory of Homosexuality for Psychoanalysis', in A. Lemma and P. E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, Hove, UK: Routledge, pp. 125–137: Primary unconscious becoming more important, as more mirroring of it: 'The discomfort of desire as made real by its social experience is here to stay. We predict that, if anything, its potential for impingement will be enhanced by the increased contingent mirroring of psychic experience made available by the panoply of electronic media, the iPhone and Android apps' (p. 135).

¹²⁷ Slade has argued that some affects are inherently difficult to mentalize, among these 'negative affects such as fear, sadness, and particularly anger'. Slade, A. (2009). 'Mentalizing the Unmentalizable: Parenting Children on the Spectrum'. *Journal of Infant, Child, and Adolescent Psychotherapy*, 8(1): 7–21, p. 8. While sexuality and aggression were given particular attention by Fonagy and Hepworth, it seems quite possible that experiences of fear will also contribute to the alien self. Fonagy and Hepworth's focus on sexuality and aggression perhaps stemmed from their salience for psychoanalytic theory. Fear may also have been subsumed under the attachment response, even though Bowlby distinguished them and their activating and terminating conditions as behavioural systems. The neglect of fear may also perhaps arise from an expectation that caregivers are able to acknowledge their child's fear more than their sexuality or aggression. If the latter assumption has been in play, it is—at least—debatable.

¹²⁸ See e.g. Freud, A. (1949). 'Aggression in Relation to Emotional Development; Normal and Pathological'. *The Psychoanalytic Study of the Child*, 3: 37–42.

And they acknowledged that many patients with severe mental health problems also report unusual or distorted concern with their bodies, which may include problems related to sexuality. However, they argued that the basis of the symptoms lies in 'the consequence of inadequate and incomplete self representation'.¹²⁹

The concern with the body that may characterize these patients is actually a secondary effect of this inadequate and incomplete self-representation: under conditions of anxiety or worry, difficulties in integrating thoughts and feelings prompt 'a movement back to a pre-reflective bodily self'.¹³⁰ Some use of this strategy is wholly ordinary, and in fact beneficial: 'the movement between physical and psychological experiences of the self is a dialectic. At times of psychological stress, we all find relief in exercise, in sexuality and other forms of regression to a pre-reflective state'.¹³¹ This movement is by no means the root of the patient's problems. As such, Fonagy and Target argued that 'classical technique, focusing upon the individual's sexual conflicts without addressing the fragility of the self structure, will rarely lead to a satisfactory therapeutic solution'.¹³² What efficacy classical technique possessed was often based, they argued, in the capacity for discussion of sexual conflicts to serve as a concrete set of metaphors for other patterns in the life of the patient. But this metaphorical use and its benefits were only accessible to patients who were already capable of reconsidering their self-representation and its functioning.

A decade later, Fonagy and Target returned to these questions in characterizing the adult experience of sexuality. This was prompted by interest in the way that consideration of sexuality had increasingly dropped out of psychoanalysis. They identified several factors implicated in this shift. One was that the object of therapeutic concern, in the aftermath of humanistic psychology, had increasingly become the interplay between self coherence and interpersonal relationships. Fonagy and Target explicitly criticized psychoanalytic thinkers such as Laplanche who emphasized sexuality at, in their view, the expense of concern with interpersonal relationships. A second factor they felt was implicated in the turn away from sexuality was that psychoanalysts felt stigmatized by the accusation that they were obsessed with sex. A third factor, Fonagy and Target argued, was that interpretations focused on sexual conflicts tended to backfire with patients with personality disorders and severe mood disorders. For instance, forms of non-mentalization could make these interpretations

¹²⁹ Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration.' *Bulletin of the Anna Freud Centre*, 17: 33–55, p. 51.

¹³⁰ Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration.' *Bulletin of the Anna Freud Centre*, 17: 33–55, p. 51. See also Fonagy, P. and Moran, G. S. (1990). 'Severe Developmental Psychopathology and Brittle Diabetes: The Motivation for Self-Injurious Behaviour.' *Bulletin of the Anna Freud Centre*, 13: 231–248: 'Emma's sole route to her dead mother was via the creation of a state of constant near-death within her own body ... The pain and discomfort she inflicted upon herself were probably only bearable because of the separation of her self representation from her representation of her physical state. As her bodily states were, to a certain extent, represented outside of her self structure, the body was available as a stage upon which the nature and functioning of the mental world could be enacted' (p. 246).

¹³¹ Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration.' *Bulletin of the Anna Freud Centre*, 17: 33–55, p. 51. See also Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418: 'Exercise in particular may not only distract depressed patients, but also reinitiate a sense of efficacy and control, lead to a libidinal reinvestment of the body, and probably have a stress-attenuating effect through its effects on the dopaminergic reward system' (p. 399). Taking this claim yet further, Fonagy has stated on Twitter that 'physical exercise is the closest we have to a panacea for mental health problems'. Accessed at: <https://twitter.com/PeterFonagy/status/1160575072068866050?s=20>.

¹³² Fonagy, P. and Target, M. (1994). 'Understanding and the Compulsion to Repeat: A Clinical Exploration.' *Bulletin of the Anna Freud Centre*, 17: 33–55, p. 51.

appear as sexual suggestions by analysts, resulting in interpersonal drama and therapeutic complications.¹³³

Fonagy and Target expressed concern that disinterest in sexuality in psychological theory put at risk attention to the alien self, present in all of us and with an especially active and disruptive role for patients with mental health problems. They speculated that caregivers are primed by evolutionary factors, often expressed in culturally specific ways, to ignore or reject signals of their child's sexual excitement. Without secondary representation and acknowledgement, sexual intentions will, and absolutely *should*, form a part of the child's alien self. Depending on the kind of non-acknowledgement a child's sexual experiences have received, this will colour the way in which they feature within the alien self: 'sexual feelings may feel too dangerous to share: they may for example be felt as too contagious and overwhelming to the relationship (unmarked mirroring), unacceptable (for example when there has been reflection of excitement and curiosity as aggression, or clear expression of disgust at the sexual body).'¹³⁴ These are the kinds of concerns associated with what Fonagy and Allison have subsequently termed the psychoanalytic unconscious. Yet states of fragmentation, incoherence, and meaninglessness may also be implicated in some of sexuality's less intentional mental states and actions: 'There is a vacuum within a part of the self, where internal reality remains nameless, sometimes dreaded, perhaps vacuous or frightening—or exciting and mysterious, as is the case in normal sexuality.'¹³⁵

There is no direct expression of sexual excitement. It is always mediated by developmental and cultural factors in how it is experienced and expressed, and by the qualities of the person or situation prompting desire. Despite this diversity, Fonagy and Target nonetheless argued that sexual excitement always entails some involvement of the alien self, even when this excitement finds expression in culturally permitted forms. Sexual pleasure comes in part from the externalization of the alien self in the excitement experienced in relation to an actual or fantasized partner.¹³⁶ The implication is that there is a fundamental incompatibility between sexual excitement and mentalizing. Holding the two together is an essentially unstable achievement at best.¹³⁷ As such, 'normal sexuality, while not pathological, mimics a form of pathology ... or is at least in the borderline spectrum.'¹³⁸ Mentalizing may help generate 'the interpersonal context for an erotically imaginative intercourse', in which a spiral of reciprocally recognized cues supports intimacy and pleasure. Nonetheless, despite the debt that these pleasures may owe to mentalizing, their urgency will also be lessened by it. We may close our eyes in sexual pleasure, Fonagy suggested, in part because close attention to the other's thoughts and feelings rather counteracts feelings of sexual excitement, since these

¹³³ Fonagy, P. (2006). 'Psychosexuality and Psychoanalysis: An Overview', in P. Fonagy, R. Krause, and M. Leuzinger-Bohleber (eds), *Identity, Gender and Sexuality: 150 Years after Freud*, London: International Psychoanalytic Association Press, pp. 1–19.

¹³⁴ Target, M. (2015). 'A Developmental Model of Sexual Excitement, Desire and Alienation', in Alessandra Lemma and Paul E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, London: Karnac Books, pp. 43–62, p. 49.

¹³⁵ Target, M. (2007). 'Is our Sexuality our Own? A Developmental Model of Sexuality Based on Early Affect Mirroring', *British Journal of Psychotherapy*, 23(4): 517–530, p. 521.

¹³⁶ Fonagy, P. (2008). 'A Genuinely Developmental Theory of Sexual Enjoyment and its Implications for Psychoanalytic Technique', *Journal of the American Psychoanalytic Association*, 56(1): 11–36: 'One's own pleasure can be experienced only when it has been placed into the other, in fantasy' (p. 24). The model is not well adapted for examining fetishes where the other will not experience pleasure. Fonagy marked this as a problem for full exploration on another day (p. 27), though without subsequently returning to the question.

¹³⁷ Target, M. (2007). 'Is our Sexuality our Own? A Developmental Model of Sexuality Based on Early Affect Mirroring', *British Journal of Psychotherapy*, 23(4): 517–530, p. 524.

¹³⁸ Fonagy, P. (2008). 'A Genuinely Developmental Theory of Sexual Enjoyment and its Implications for Psychoanalytic Technique', *Journal of the American Psychoanalytic Association*, 56(1): 11–36, p. 19.

are premised on localized non-mentalizing in the externalization of the alien self.¹³⁹ Target has added that sexual pleasures may be helped along by other modes of non-mentalizing. We may experience pleasure as a demonstration of the strength or quality of the relationship (psychic equivalence). A partner's pleasure may be experienced as concretely affirming their love (teleological mode). Or we may take pleasure from feelings of being merged with the other and safe from the world (pretend mode).¹⁴⁰

The non-mentalizing of externalization of the alien self in sexuality may have an evolutionarily adaptive function. Not only does it direct sexuality away from the family during adolescence, helping to avoid incest. Fonagy and Target also offered their suspicion that the potential for localized heedlessness and recklessness in sexual life is adaptive for reproductive success, even if it may be very inconvenient in other ways.¹⁴¹ They proposed that, at an evolutionary level, sexual desire is in tension with the evolutionary benefits of stability and interpersonal understanding. Likewise, at the level of the individual, Fonagy and Target proposed that there is a trade-off between sexual excitement and the other benefits of relationships. The better integrated self-representation afforded by the relationship as a secure base and safe haven quietens the demands of the alien self, and with it the intensity of sexual excitement. The partner is partly internalized through preconscious identification, as a shared sense of 'we', which replaces the sense of mystery and excitement with familiarity and security.¹⁴² Though many relationships may be able to toggle backwards and forwards between the two, Fonagy and Target suggest that the relative incompatibility of sexuality and mentalizing is one of several reasons that desire tends to fade over time in long-term relationships. When this happens, sexual excitement may still be evoked by other people, they argued, to the extent that these individuals elicit aspects of the alien self that have not found purchase and integration within the present relationship.¹⁴³

On the basis of this account, Fonagy and Target argued that sexuality should neither be the central focus of therapy nor should it be ignored or treated only as a metaphor. When sexual excitement appears in a therapeutic context, the therapist is advised to offer 'recognition of the feelings, minor resonance without reciprocal excitement, without denial or distortion'.¹⁴⁴ Sexuality is part of the alien self for a reason: its expression is often inappropriate and difficult to reconcile with mentalizing. Yet sexuality is relevant in therapeutic contexts, because it can provide important signals of the activity of the alien self. Sexuality may be used as 'a vehicle for carrying repudiated aspects of the self such as aggression, envy, grandiosity,

¹³⁹ Ibid. 26. However, there may be gender differences here: 'I believe that while male sexual enjoyment culminates in the full externalization of the self into the object and its unconsciously fantasied control therein, female sexual arousal begins with an intersubjective identification with the partner and becomes increasingly "private" and inwardly turning as excitement mounts. In both cases, intersubjectivity is critical to fulfillment, but while male excitement moves toward seeing the split-off self as the other, the vector or focus of female excitement is an increasingly direct experience of a self uncontaminated by incongruity, assuming that a previous successful projection has taken place' (p. 33).

¹⁴⁰ Target, M. (2015). 'A Developmental Model of Sexual Excitement, Desire and Alienation', in Alessandra Lemma and Paul E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, London: Karnac Books, pp. 43–62, p. 46–47.

¹⁴¹ Fonagy, P. (2008). 'A Genuinely Developmental Theory of Sexual Enjoyment and its Implications for Psychoanalytic Technique'. *Journal of the American Psychoanalytic Association*, 56(1): 11–36, p. 24.

¹⁴² Target, M. (2015). 'A Developmental Model of Sexual Excitement, Desire and Alienation', in Alessandra Lemma and Paul E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, London: Karnac Books, pp. 43–62, pp. 51–52.

¹⁴³ Fonagy, P. (2008). 'A Genuinely Developmental Theory of Sexual Enjoyment and its Implications for Psychoanalytic Technique'. *Journal of the American Psychoanalytic Association*, 56(1): 11–36, p. 25.

¹⁴⁴ Target, M. (2007). 'Is our Sexuality our Own? A Developmental Model of Sexuality Based on Early Affect Mirroring'. *British Journal of Psychotherapy*, 23(4): 517–530, p. 527.

derogation, contempt, or plain selfishness.¹⁴⁵ Put another way, the most potent metaphors rub both ways, with true aspects as well as pretend aspects at once. Even if sexuality is in part a metaphor for other troubles, this is grounded in the powerful alien self aspects of sexuality:

Sexuality is not at root conflictual; rather, conflicts come to be expressed via the sexual metaphor. It is this psychic flypaper quality that makes psychosexuality such a key part of understanding our patients. Many truly painful conflicts are sexual, not because they are rooted there but because the otherness quality of sexuality frames the conflict as being external. As the psychosexual expresses, and does not disguise, the relational, frequently the only genuine route to understanding relational issues is through psychosexual experience.¹⁴⁶

When therapists direct attention away from sexuality, this leaves the patient without support for making sense of and modulating feelings that may be sexual or that may be framed in terms of sexuality. Fonagy and colleagues hold that sexuality should be recognized by therapists, but so as to focus on supporting the patient's capacity to mentalize related thoughts and feelings. Discussion of sexual matters may be inhibited in short-term therapeutic work by the patient's shame or worries. And over time, the experience of the therapist as a secure base can be expected to inhibit sexual excitement. But at the point that the therapeutic relationship is evoking the alien self, and remains exciting and mysterious, sexuality is likely to be an especially important theme.

Having considered sexuality, we can now turn to a second, parallel contributor to the alien self in the work of Fonagy and colleagues. This is aggressive intentions. Whereas Fonagy and colleagues describe sexuality in terms that suggest that externalization of the alien self is generally involved, their stance appears to be less ambitious in relation to aggression, seeing externalization as an important cause but not implicated in aggression in all its forms.¹⁴⁷ By contrast, the multiple causal factors that contribute to aggression are acknowledged especially in writing for developmental psychology audiences, in the context of a discipline in which unitary causal lines are not regarded as persuasive.¹⁴⁸ For instance, Fonagy has recently highlighted the importance of economic deprivation for the prevalence of conduct problems, without implying that this relationship is fully mediated by externalization of the alien self.¹⁴⁹ Nonetheless, the writings of Fonagy and colleagues suggest that externalization of the alien self may have a significant role to play in the initiation and maintenance of violence.

In classical psychoanalysis, a child's aggressive wishes—like their sexual desires—cannot be fully realized and are prohibited by both caregivers and wider society. Again, for Freud, the conflict between unconscious aggressive intentions and conscious recognition of the inappropriateness of aggression was anticipated to be implicated in the symptoms of patients.¹⁵⁰ Subsequent psychoanalytic theorists such as Kernberg built on this position to

¹⁴⁵ Fonagy, P. (2009). 'Foreword', in C. Clulow (ed.), *Sex, Attachment and Couple Psychotherapy*, London: Karnac Books, p. xx.

¹⁴⁶ Fonagy, P. (2008). 'A Genuinely Developmental Theory of Sexual Enjoyment and its Implications for Psychoanalytic Technique', *Journal of the American Psychoanalytic Association*, 56(1): 11–36, p. 28.

¹⁴⁷ Though admittedly what appears as a theoretical difference may be an artefact of genre. Fonagy and colleagues barely mention sexuality outside of their writings for psychoanalytic audiences.

¹⁴⁸ Cicchetti, D. and Rogosch, F. A. (1996). 'Equifinality and Multifinality in Developmental Psychopathology', *Development & Psychopathology*, 8: 597–600.

¹⁴⁹ Fonagy, P. (2019). 'The Future Prospects of Mentalization Based Therapies', 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November.

¹⁵⁰ See Freud, A. (1949). 'Notes on Aggression', *Bulletin of the Menninger Clinic*, 13(5): 143–151.

argue that excessive destructive and envious intentions, combined with a caregiving environment that was not able to offer containment of these responses, predispose BPD.¹⁵¹ They described by Fonagy, Moran, and Target, the authors set themselves against this account. They argued that aggression might better be seen as a defence against particular thoughts and feelings.

They described Moran's treatment of a patient, David. On the basis of the concept of mentalizing theory, they argued that his 'unprovoked, excessively aggressive acts could now be seen to be aimed at objects whom he experienced as viewing him in a negative way. It seemed as if, through attacking these individuals, he could temporarily rid himself of bad reflections of himself ... It was preferable for David to strip his mind almost bare of mental content than to be exposed to the terrible power of the shaming gaze.'¹⁵² This account allowed the authors to pinpoint the conditions under which David would show aggression: 'his aggression was activated on the numerous occasions when his sense of self required reinforcement and his identity was in need of bolstering'.¹⁵³ Though the concept would only be offered some years later, we can see in this characterization the proposal that aggression represented a means for David to externalize an alien self associated with shame. This defence was interpreted as reflecting David's difficulties in conceiving of and reflecting on his mental states and those of others. However, the defence was achieved precisely by reinforcing and solidifying this state of affairs.

Given that all humans have an alien self, rather than solely ask how violent behaviour develops, Fonagy and Target's perspective suggested that another focus should be on how a failure to inhibit violent behaviour develops.¹⁵⁴ Fonagy and Target proposed that violence did not reflect a single developmental trajectory: multiple developmental trajectories might converge on deficits in mentalizing combined with thoughts or feelings experienced as unbearable and as requiring externalization. Pervasively aggressive behaviour was therefore interpreted by Fonagy and Target as reflecting, and reinforcing, 'a wish to attack thoughts, in oneself or in another'.¹⁵⁵ For instance, they described self-harm as an externalization into the pre-reflexive body of intolerable thoughts or feelings that would otherwise occur in the patient's own mind. The ritualized and structured form of an act of self-harm offers a way to access a certain kind of regularized feeling, and therefore provide a certain sense of clarity.

¹⁵¹ Fonagy's impression of Bion and Kernberg's positions is set out in Fonagy, P. and Higgitt, A. (1990). 'A Developmental Perspective on Borderline Personality Disorder'. *Revue Internationale de Psychopathologie*, 1: 125–159, p. 142.

¹⁵² Fonagy, P., Moran, G. S., and Target, M. (1992). 'Aggression and the Psychological Self'. *Bulletin of the Anna Freud Centre*, 15: 269–284, pp. 279–80.

¹⁵³ *Ibid.* 281.

¹⁵⁴ This was well aligned with the emergent perspective of developmental psychopathology e.g. Rutter, M. and Sroufe, L. A. (2000). 'Developmental Psychopathology: Concepts and Challenges'. *Development and Psychopathology*, 12(3): 265–296: 'The search needs to be for factors that fail to inhibit aggressive-disruptive behaviour, as well as for factors that foster it' (p. 276). More recently, see Patalay, P., Fink, E., Fonagy, P., and Deighton, J. (2016). 'Unpacking the Associations between Heterogeneous Externalising Symptom Development and Academic Attainment in Middle Childhood'. *European Child & Adolescent Psychiatry*, 25(5): 493–500.

¹⁵⁵ Fonagy, P. and Target, M. (1995). 'Understanding the Violent Patient: The Use of the Body and the Role of the Father'. *The International Journal of Psychoanalysis*, 76: 487–501, p. 489. The authors also offered some unpersuasive reflections on gender differences in violence. *Ibid.*: 'Why do aggressive men more often direct their hostility towards others, while self-mutilation is more common in women? We believe that both forms of violence suggest an attempt to be rid of an intolerable fantasy of the thoughts in somebody's mind, originally the thoughts of a parent. The gender imbalance may then reflect a wish to attack the thinking of the same-sex parent (with whom identification is potentially more painful and inescapable). For both girls and boys, the mother's thoughts about the child have generally been intersubjectively experienced earlier, and are probably represented as within the child's mind. The father's thinking is, we suggest, represented in both sexes as external. The intolerable mental presence of the same-sex parent would then be felt to be inside the woman's mind, but outside the man' (p. 498).

And this is achieved in such a way as to direct attention to the body, and away from mental states. Likewise, Fonagy theorized that male violence against women represents an attack on thinking, this time through the body of the other. He was interested by the way that, in his experience, perpetrators reported that control over and intimidation of women offered them a feeling of personal coherence.¹⁵⁶

In a paper from 2003, Fonagy characterized violence that functions to support the feeling of personal coherence as 'representational violence', because the object of the violence is used to 'represent' the aspects of the alien self that are refused mentalization.¹⁵⁷ He contrasted this with a second basis for violence, 'violence in the negative'. In this form, states of fragmentation, incoherence, and meaninglessness from the primary unconscious have become associated with specific other people or groups. These then have to be 'destroyed to protect the mind from states which are experienced as out of control'.¹⁵⁸ When some capacity for mentalization is available, the limited responsibility of the other for these states can usually be acknowledged. When mentalization is offline, and especially when teleological mode is active, physical force may be regarded as necessary in order to achieve control and change in the situation. Fonagy summarized: 'representational violence' is aimed at an object in its function as a signifier for mental states that are refused to be conceived or considered as thoughts or feelings; 'violence in the negative' is aimed at a signifier in order to prevent its expression of states of fragmentation, incoherence, and meaninglessness.

However, in another article of the same year, Fonagy argued that not all violence fits into this scheme, and that psychological theory 'must take proper account of the existence of positive, survival-oriented aggression and also of aggression that is a genuine protest against hardship in life'.¹⁵⁹ He argued that violence should not be denigrated: evolutionary pressures will have required humans to be capable of both nonviolent group life and of enacting violence under situations of threat, proposing that mentalization may have evolved as a capacity in part to help sustain nonviolence in social contexts where we feel recognized, while prompting non-mentalizing and violent responses in environments of nonrecognition.¹⁶⁰ This perspective has significant implications for prevention (see Chapter 9). Similarly, the idea of social acknowledgement as a prompt, primed by evolution, for mentalizing rather than violence may help explain why 'restorative justice' approaches, in which an offender and victim of crime interact face to face and offer mutual acknowledgement, have been found to be especially powerful at reducing violent crime, compared with other less serious forms of crime.¹⁶¹ Restorative justice approaches differ markedly from the usual approach of criminal justice systems, which institutionalize teleological mode thinking in demanding concrete retribution for crime, even in the face of evidence that this appears to make future crime more, rather than less, likely.

¹⁵⁶ Fonagy, P. (1999). 'Male Perpetrators of Violence against Women: An Attachment Theory Perspective.' *Journal of Applied Psychoanalytic Studies*, 1(1): 7–27.

¹⁵⁷ Fonagy, P. (2003). 'The Violence in our Schools: What can a Psychoanalytically Informed Approach Contribute?'. *Journal of Applied Psychoanalytic Studies*, 5(2): 223–238. Representational violence: 'The first kind of violence arises because of the overriding need for coherence in the experience of an agentive self... For its effectiveness, it depends on the experience of having created an intentional state in the victim that enhances the strength and coherence of the self by externalization.' (pp. 232–233).

¹⁵⁸ Ibid. 235.

¹⁵⁹ Fonagy, P. (2003). 'Towards a Developmental Understanding of Violence', *British Journal of Psychiatry*, 183: 190–192, p. 190.

¹⁶⁰ Ibid. 191.

¹⁶¹ Fonagy, P. (2012). 'The Neuroscience of Prevention'. *Journal of the Royal Society of Medicine*, 105(3): 97–100, p. 99.

In recent years, and in line with other perspectives in developmental psychopathology,¹⁶² Fonagy has increasingly drawn distinctions between mild aggressive problems and more intractable forms of antisocial personality disorder, while continuing to regard deficits in mentalizing as relevant to both. Various developmental pathways may converge in contributing to the inhibition of mentalizing, which can contribute to the externalization of the alien self, reduced feelings of personal agency, hostile attributions regarding others, and a low threshold for violent actions.¹⁶³ In reflecting on these developmental pathways, Fonagy has acknowledged the likelihood that 'some particularly intractable conduct problems have a genetic basis'.¹⁶⁴ He has not held out expectation for an 'aggressive genotype'. Rather, he has been more persuaded by the idea that there may be both genetic and environmental contributions to attentional control and affect regulation, which help sustain mentalizing. This conclusion is supported by findings that 90% of children with a diagnosed conduct disorder also have another mental health problem—for instance, with attentional problems 10 times higher than in the general population.¹⁶⁵

What especially distinguishes antisocial personality disorder, for Bateman and Fonagy, are feelings of shame in the alien self, which are externalized through the domination of other people. This externalization is prompted by psychic equivalence, where shame threatens to dominate and define reality.¹⁶⁶ The externalization of the alien self functions through the denigration and intimidation of other people, which holds at bay questions these others or the self might pose about the individual's worth. Any potential or perceived slight is met with hostility, perhaps supported by teleological mode in which physical retribution seem the only real response.¹⁶⁷ Like the other forms of non-mentalizing, the externalization of the alien self presents specific obstacles to the modulated doubt that helps sustain the inquisitive stance and consideration of alternate perspectives. Externalization of the alien self seeks to control the other's behaviour, motivations, and experience, and in this way to avoid relevant feelings and thoughts that might otherwise disrupt a brittle self-coherence. The question of how to conceptualize whether such strategies are 'adaptive' will be pursued further in the next chapter. By the end of the 2000s, this account of adaptation had come to rather eclipse the idea of the alien self, which remained 'on the books' but without any scientific operationalization and poorly integrated with subsequent theoretical developments. This may be attributed to the essentially psychoanalytic underpinnings of the concept of externalization of the alien self, which has meant that the concept travels less easily, and is less

¹⁶² E.g. Mesman, J., Stoel, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Juffer, F., Koot, H. M., and Alink, L. R. (2009). 'Predicting Growth Curves of Early Childhood Externalizing Problems: Differential Susceptibility of Children with Difficult Temperament'. *Journal of Abnormal Child Psychology*, 37(5): 625–636.

¹⁶³ Taubner, S., Gablonski, T.-C., and Fonagy, P. (2019). 'Conduct Disorder', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 301–321.

¹⁶⁴ Fonagy, P. (2012). 'The Neuroscience of Prevention'. *Journal of the Royal Society of Medicine*, 105(3): 97–100, p. 98.

¹⁶⁵ Pilling, S. and Fonagy, P. (2012). 'Developing Clinical Guidelines for Children and Adolescents: Experience from the National Institute for Health and Clinical Excellence', in P. Sturmey and M. Hersen (eds.), *Handbook of Evidence-Based Practice in Clinical Psychology*, Volume 1, Child and Adolescent Disorders, New York: Wiley, pp. 73–102, p. 74.

¹⁶⁶ Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). 'The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence'. *Psychoanalytic Inquiry*, 23(3): 412–459, p. 446.

¹⁶⁷ Bateman, A. and Fonagy, P. (2012). 'Antisocial Personality Disorder', in A. W. Bateman and P. Fonagy (eds.), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 289–308; Bateman, A., O'Connell, J., Lorenzini, N., Gardner, T., and Fonagy, P. (2016). 'A Randomised Controlled Trial of Mentalization-Based Treatment versus Structured Clinical Management for Patients With Comorbid Borderline Personality Disorder and Antisocial Personality Disorder'. *BMC Psychiatry*, 16(1): 304.

intelligible to the scientific and policy audiences of the more recent work of Fonagy and colleagues.¹⁶⁸ Nonetheless, Fonagy and colleagues' recent interest in wider social systems and institutionalized non-mentalizing have drawn attention to experiences of social alienation (see Chapter 9). It may be that the alien self will receive renewed interest, through this link with alienating social systems.

¹⁶⁸ It may be that the reduced prominence of externalization of the alien self can also be attributed to the changing membership of Fonagy's closest collaborators, with Target working increasingly less closely with Fonagy in recent years.

Adaptation and mental health

Introduction

Dominant models of mental illness often treat it as a discrete pathology, distinct from ordinary human psychology. One of the more subversive aspects of psychoanalysis was the desire to look at typical and atypical psychological phenomena together, and examine variation between them in terms of common mechanisms.¹ This included speculations about neurological mechanisms relevant to both the makings of ordinary life and the formation of symptoms. Dreams, jokes, and the slips of everyday life were included by Freud within the same framework as obsessional symptoms, somatic disorders, and psychosis:

The theoretical importance of this conformity between dreams and symptoms is illuminating. Since dreams are not pathological phenomena, the fact shows that the mental mechanisms which produce the symptoms of illness are equally present in normal mental life, that the same uniform law embraces both the normal and the abnormal and that the findings of research into neurotics or psychotics cannot be without significance for our understanding of the healthy mind.²

There was a price to this strategy for psychoanalytic theory. Many of the concepts used to link typical and atypical phenomena—such as ‘defence’—were forced to become over-absorptive of different meanings, to the point that their meaning became lost. Yet at the same time, Freud intended that this approach would reduce the stigma attached to mental illness and support preventative work with people who were not yet unwell.

Very much the same situation has played out for the concept of ‘mentalizing’. This concept likewise was treated by Fonagy as relevant to typical and atypical mental development, of transdiagnostic significance, and not wholly conscious. Though Fonagy and colleagues would regard their claims about neurology to be less speculative than Freud’s, these claims have likewise addressed mechanisms relevant to both ordinary life and the formation of symptoms. Like Freud, the simultaneous attention to typical and atypical development by Fonagy and colleagues has been pursued in part as a way to combat the ‘unfounded, crass, insuperable prejudice’ associated with mental illness.³ There have been theoretical gains associated with this approach. Insight into successful mentalizing offered ideas for conceptualizing mental ill health; and insight into forms of non-mentalizing offered ideas for conceptualizing the structure of the self, the nature of perceptual experience, and the specific

¹ Freud, S. ([1901] 2001). *The Psychopathology of Everyday Life*, in James Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 6, London: Vintage.

² Freud, S. ([1923] 2001). ‘Two Encyclopaedia Articles’, in James Strachey (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Volume 18, London: Vintage, pp. 233–260, p. 241.

³ Luyten, P. and Fonagy, P. (2012). ‘The Multidimensional Construct of Mentalization and its Relevance to Understanding Borderline Personality Disorder’, in A. Fotopoulou, D. Pfaff, and M. A. Conway (eds), *From the Couch to the Lab: Trends in Psychodynamic Neuroscience*, Oxford: Oxford University Press, pp. 405–427, p. 406.

qualities of psychological insight. However, like several of Freud's concepts, along the way, uses of the idea of mentalization expanded, and expanded.

This prompted the work of Chapters 4 and 5 to attempt to pin down more precisely the definition and constituent elements of mentalization and non-mentalization, and to ask whether they are simple opposites. It also prompted effort in the previous chapter to draw out the complex picture of how Fonagy and colleagues have discussed the role of unconscious processes in everyday life, including in experiences of sexuality and aggression. In this chapter, we turn to consideration of Fonagy and colleagues' conceptualization of the nature of mental illness, predominantly in papers written for researchers in the field of developmental psychopathology, though with some eye to clinical audiences. We start with criticisms of Fonagy and Target's work from the 1990s, raised by allies such as Jeremy Holmes and Otto Kernberg who alleged that mentalization risked becoming a disempowering, deficit-focused model of mental ill health. One resource available to Fonagy and colleagues in responding to this concern was attachment theory, which had developed a theory of attachment strategies as evolutionarily primed responses to adverse conditions. We will survey the reflections Fonagy and colleagues have offered on three attachment theorists—Mary Main, Pat Crittenden, and Jay Belsky. The synthesis proposed by Fonagy and colleagues will then be described. The position of Fonagy and colleagues gives prominence to attachment relationships in calibrating an individual's sense of vigilance or trust in the perspectives of others, as the precondition of learning from them. In recent years, this concern with vigilance, trust, and learning has been central to the theoretical work of Fonagy and colleagues, laying the ground for very important advances in specifying the role of mentalization in both ordinary life and the alleviation of mental ill health.

Adaptation and attachment

Through the 1990s and at the start of the 2000s, Fonagy, Target, and collaborators had a tendency to denigrate non-mentalizing. They described non-mentalizing as an 'unholy perversion of reflection',⁴ and as a 'stunted or twisted' version of mentalization.⁵ The ambition in these passages appears to have been to offer lively description of the way in which forms of non-mentalizing seem to utilize many of the same component elements as mentalization—for instance, that pretend mode utilizes the imaginative capacity to conceive of mental states that is required, too, for mentalization. However, the language is also stigmatizing, perhaps an unintended reflection of the authors' clinical frustration with recalcitrant cases. In the mid-2000s, important criticisms were raised of the assumptions about mentalization that such language seemed to express. Jeremy Holmes and Otto Kernberg are psychoanalytic psychiatrists, and friends and allies of Fonagy and Target. Both were early advocates for mentalization in the 1990s,⁶ and have subsequently co-authored work with Fonagy.⁷ Yet,

⁴ Bach, S., Mayes, L., Alvarez, A. and Fonagy, F. (2000). 'Panel 1: Definition of the Self'. *Journal of Infant, Child & Adolescent Psychotherapy*, 1(3): 5–24, p. 21.

⁵ Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873, p. 859.

⁶ Kernberg, O. F. (1997). 'The Nature of Interpretation: Intersubjectivity and the Third Position'. *The American Journal of Psychoanalysis*, 57(4): 297–312; Holmes, J. (1998). 'The Changing Aims of Psychoanalytic Psychotherapy: An Integrative Perspective'. *The International Journal of Psychoanalysis*, 79: 227–240.

⁷ E.g. Shedler, J., Beck, A. T., Fonagy, P., Gabbard, G. O., Kernberg, O., Michels, R., and Westen, D. (2011). 'Response to Skodol Letter'. *American Journal of Psychiatry*, 168(1): 97–98; Chiesa, M., Fonagy, P., Holmes, J., and Drahorad, C. (2004). 'Residential versus Community Treatment of Personality Disorders: A Comparative Study of Three Treatment Programs'. *American Journal of Psychiatry*, 161(8): 1463–1470. See also Hepworth,

they criticized Fonagy and Target for tending to characterize non-mentalizing as a 'deficit' or 'impairment', without considering whether it may have benefits or adaptive aspects.⁸ This was out of keeping with a growing critique of deficit thinking in mental health in the 2000s: it was held that mental health theory and practice should recognize the strengths of users of services, acknowledge their agency beyond that 'granted' by professionals, and differentiate between short-term and long-term advantages and disadvantages.⁹ For instance, substance use might be reframed as, in certain contexts, a form of self-medication against other difficulties.¹⁰

Fonagy and colleagues accepted the criticism of Holmes and Kernberg as partially justified. There was a tendency towards deficit-thinking in their theorizing.¹¹ The criticism was not wholly justified, however. Fonagy and Target had conceptualized symptoms of mental ill health as intelligible responses to adverse situations. One of the characteristic features of the Anna Freudian tradition, within which Fonagy and Target had trained, was a concern with 'lines of development which lead from the individual's state of infantile immaturity and dependence to the gradual mastery of his own body and its functions, to adaptation to the object world, reality and the social community' (pp. 47–53).¹² Anna Freud identified this adaptation with the ego.¹³ She opposed the ego to the id, conceptualized as continually seeking 'fulfilment of drive derivatives which are unacceptable to the environment, bringing with it the threat of punishment or loss of love, and, instead of serving adaptation, disturbing it.'¹⁴ Defence mechanisms were regarded by Anna Freud as compromises between the ego and the id, representing the best available adaptation of the child to his or her environment. This adaptation would be shaped by the extent of supports available in that environment, and the extent of effective dominance of the ego over the id. In this subtle perspective, depending

M. (2019). "Peripheral Visions": A Conference Celebrating the Contribution of Jeremy Holmes and 20 Years of Psychoanalytic Studies at Exeter University, UK. *British Journal of Psychotherapy*, 35(2): 215–216.

⁸ Holmes, J. (2005). 'Notes on Mentalizing—Old Hat, or New Wine?', *British Journal of Psychotherapy*, 22(2): 179–198; Kernberg, O. F., Yeomans, F. E., Clarkin, J. F. and Levy, K. N. (2008). 'Transference Focused Psychotherapy: Overview and Update'. *The International Journal of Psychoanalysis*, 89: 601–620.

⁹ E.g. Blundo, R. (2001). 'Learning Strengths-Based Practice: Challenging our Personal and Professional Frames'. *Families in Society*, 82(3): 296–304; Crossley, N. (2004). 'Not being Mentally Ill: Social Movements, System Survivors and the Oppositional Habitus'. *Anthropology & Medicine*, 11(2): 161–180.

¹⁰ Khantzian, E. J. (1997). 'The Self-Medication Hypothesis of Substance Use Disorders: A Reconsideration and Recent Applications'. *Harvard Review of Psychiatry*, 4(5): 231–244.

¹¹ See e.g. Allison, E. and Fonagy, P. (2016). 'When is Truth Relevant?'. *Psychoanalytic Quarterly*, 85(2): 275–303, p. 286. They now acknowledge 'when we think about mental ill health, particularly in children and young people, inorganic metaphors of damage or even plasticity may be less helpful than ways of thinking that recognize the human capacity for growth and adaptation in conditions. Of course, this capacity is not infinite and can be constrained by both biological and environmental factors, but it is at times grossly underestimated both by those who tend to think of disorder in terms of irreparable damage and by well-meaning professionals who believe that recovery occurs only as a consequence of intervention', Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

¹² Freud, A. (1962). 'Assessment of Childhood Disturbances'. *The Psychoanalytic Study of the Child*, 17(1): 149–158, p. 152–3. See also Joffe, W. G. and Sandler, J. (1979). 'Adaptation and Individuation'. *Bulletin of the Anna Freud Centre*, 2(3): 127–161; Sandler, A.-M. (2012). 'Anna Freud's Influence on Contemporary Thinking about the Child', in N. Malberg and J. Raphael-Leff (eds), *The Anna Freud Tradition: Lines of Development*, London: Karnac Books; Midgley, N. (2012). *Reading Anna Freud*, London: Routledge.

¹³ The work of Anna Freud and Heinz Hartmann would together prove influential for the central role of the concept of 'adaptation' within subsequent ego psychology. See e.g. Rapaport, D. and Gill, M. M. (1959). 'The Points of View and Assumptions of Metapsychology'. *The International Journal of Psychoanalysis*, 40: 153–162. It is commonly pointed out that 'adaptation' receives special attention in the work of these central European Jewish émigrées, who had to acclimatize to America and England. Two generations later, there seem relevant analogies to the case of Fonagy.

¹⁴ Freud, A. (1981). 'Insight: Its Presence and Absence as a Factor in Normal Development'. *The Psychoanalytic Study of the Child*, 36(1): 241–249, p. 243.

on development and context, exactly the same behaviour could be in the service of the ego and adaptation, or represent defence and the compromise of ego and id:

A five-year-old may ask his mother for an extra cardigan 'so as not to catch cold'; or a six-year-old may anxiously wash his hands before dinner since 'there might be germs'. Should that be understood as early compliance with adult rules, i.e. a welcome sign of adaptation, or, rather, as an ominous sign of hypochondriacal tendencies and of pathogenic defence mechanisms used in the struggle against anal messing? Obviously, only an approximate timetable for normal advance towards compliance with hygiene and cleanliness can supply the answer.¹⁵

In the 1990s, when Fonagy and colleagues were working on a manual for child psychoanalysis at the Anna Freud Centre, they adopted an aligned perspective, arguing that 'the child needs his defences and has good reason for erecting them'.¹⁶ Speaking on BBC television, Fonagy criticized therapeutic approaches that do not take into account the difficult feelings or the problems in a child's environment against which defences may have been elaborated: 'if you are systematically undermining that very fragile, that very vulnerable sense of who that child is you could end up in the situation where the child becomes really very much more depressed and hopeless and helpless. My concern about it is that you are destroying something in the child that is the child's own. However distorted and however maladaptive it is, it is the child's own'.¹⁷

The concept of 'adaptation' would be a piece of the inheritance of Fonagy and his collaborators from the Anna Freudian tradition.¹⁸ Yet a problem with the Anna Freud's concept of 'adaptation' was its multiple meanings, a problem raised with Freud late in her life by Joseph

¹⁵ Freud, A. (1983). 'The Past Revisited'. *Bulletin of the Anna Freud Centre*, 6(1): 107–113, p. 111.

¹⁶ Fonagy, P., Edgumbe, R., Target, M., and Miller, J. (1999). *Contemporary Psychodynamic Child Therapy: Theory and Technique*, London: Anna Freud Centre and University College London: 'Early in treatment, it is usually safer to begin with verbalisation of affect until the child's emotional balance is clear enough for the analyst to predict how he might react to interpretation of unconscious wishes. Even so, it is also usually more readily acceptable to the child to interpret defences before the wishes or impulses against which the child is defending. Bypassing the defences with a direct interpretation of an unconscious wish ignores the fact that the child needs his defences and has good reason for erecting them' (p. 103). See also Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38: 'For these children, the possibility of relinquishing their single, fixed, concrete model of self-other relatedness triggers terror and confusion. Therapists more profitably point out the advantages of not changing, remarking on the price children would pay if they were to give up their maladaptive, but often life-saving, defences and self-other internal models' (p. 31).

¹⁷ BBC2 (2001). 'Taming the Problem Child, 9 p.m. Horizon, Thursday, 8 March. Accessed at: http://www.bbc.co.uk/science/horizon/2000/problemchild_transcript.shtml.

¹⁸ Target, M. (2018). '20/20 Hindsight: A 25-year Programme at the Anna Freud Centre of Efficacy and Effectiveness Research on Child Psychoanalytic Psychotherapy', *Psychotherapy Research*, 28(1): 30–46: 'Our effort in this direction, the HCAM [Hampstead Child Adaptation Measure], distinguished 14 domains of normal and abnormal development which could be tracked cross-sectionally and longitudinally, going well beyond symptoms to cover things like the child's capacity to have relationships, to play, learn, look after himself physically, regulate moods, manage stress, and so on' (p. 43). To take another example, a scale for the extent to which an individual's behaviour appears adaptive would be developed by Steele, H., Steele, M., and Kriss, A. (2009). *The Friends and Family Interview (FFI) Coding Guidelines*. Unpublished manuscript: 'This scale refers specifically to responses to the question asking what the respondent does when distressed or upset. An adaptive strategy may involve seeking comfort from others (e.g. parents, friends, or siblings), engaging in a favorite activity that relieves their unhappiness (e.g. listening to music, walking the dog), or simply thinking things through.' See also Chow, C. C., Nolte, T., Cohen, D., Fearon, R. P., and Shmueli-Goetz, Y. (2017). 'Reflective Functioning and Adolescent Psychological Adaptation: The Validity of the Reflective Functioning Scale-Adolescent Version'. *Psychoanalytic Psychology*, 34(4): 404–413.

Sandler, and which she seemed to accept.¹⁹ What may be adaptive in the short run may not be adaptive in the long run, and vice versa. What may be adaptive for a species may not be adaptive for an individual. Furthermore, the term 'adapt' could mean simply the extent to which an individual can respond to a challenge, or the extent to which they can thrive in the face of it.²⁰ There can also be a subtly normative aspect, in which adaptation to an environment may mean capitulation to its norms and existing power relations.²¹ By the time Fonagy was developing his ideas in the 1980s and 1990s, he was in discussion with psychoanalysts such as Sandler who were acknowledging these various meanings of the concept of adaptation, and the threat that the concept might smuggle unacknowledged assumptions about what counts as the good life. But psychoanalytic discourse frequently lost track of the distinctions. And there was no strong theory available for how individual responsiveness and thriving might relate to the evolutionary level of adaptation.

In this context, Fonagy and colleagues took particular interest in developments in attachment theory in conceptualizing adaptation, defences, and mental health and ill health. These developments had been prompted by an influential, but unpublished, conference paper presented by Mary Ainsworth at the International Conference on Infant Studies in April 1984.²² In this paper, Ainsworth expressed concern that the term 'adaptation' represented a twig-thicket of different meanings; attention to these different meanings offers clarification on the relationship between mentalizing and 'adaptation'. Like Sandler, Ainsworth acknowledged that 'adaptation' could refer to processes at a species level, in identifying a behavioural system or trait as contributing to survival or reproduction. Secondly, 'adaptation' could refer to an individual level, identifying a behaviour or trait as responsive to the available rewards and punishments of the immediate environment.²³ Ainsworth also observed a third meaning of the concept: 'In the developmental mental health sense the focus is on how individual differences in development, and on evaluation of how well or how poorly such development equips the individual to cope with the impact of the environment in which he lives.'²⁴ What distinguished this third meaning of the concept from the second was that an evaluation was entailed. The second meaning was merely an acknowledgement that an individual may

¹⁹ Sandler, J. and Freud, A. (1981). 'Discussions in the Hampstead Index on "The Ego and the Mechanisms of Defence": V. The Mechanisms of Defence, Part 2'. *Bulletin of the Anna Freud Centre*, 4(4): 231–277; Sandler, J. and Freud, A. (1982). 'Discussions in the Hampstead Index on "The Ego and the Mechanisms of Defence": VIII. Denial in Word and Act'. *Bulletin of the Anna Freud Centre*, 5(3): 175–187. See also Joffe, W. G. and Sandler, J. (1968). 'Comments on the Psychoanalytic Psychology of Adaptation, with Special Reference to the Role of Affects and the Representational'. *The International Journal of Psychoanalysis*, 49: 445–454. Early concerns about the term 'adaptation' were raised by Hartmann, one of the popularizers of the concept within the psychoanalytic community. See Hartmann, H. (1939). 'Psycho-Analysis and the Concept of Health'. *The International Journal of Psychoanalysis*, 20: 308–321.

²⁰ Winnicott, in turn, used the term 'adaption' in a different, slightly non-standard way, developing further the idea of 'responding'. His predominant use of the term characterized the way that a caregiver moulds his or her body and interactions to fit around those of the child. Winnicott, D. W. ([1960] 1965). *The Maturation Process and the Facilitating Environment*, New York: International Universities Press.

²¹ See also Lee, N. N. (2014). 'Sublimated or Castrated Psychoanalysis? Adorno's Critique of the Revisionist Psychoanalysis: An Introduction to "The Revisionist Psychoanalysis"'. *Philosophy & Social Criticism*, 40(3): 309–338.

²² Ainsworth, M. (1984). 'Attachment, Adaptation and Continuity'. Paper presented at the International Conference on Infant Studies, April, John Bowlby Archive, Wellcome Collections, PP/Bow/J.1/57. See Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press, Chapter 2.

²³ *Ibid.*: 'In the phylogenetic or evolutionary sense adaptation implies that in the course of natural selection those behaviours that yield survival advantage in the environment in which the evolutionary changes are taking place become part of the behavioural repertoire characteristic of a species ... In the ontogenetic sense adaptation refers to the process through which an organism adjusts to its environment in the course of development.'

²⁴ *Ibid.*

'adapt' to their circumstances. Ainsworth's third meaning was to identify 'adaptation' as the capacity to thrive in the long term within those circumstances.

Ainsworth's argument was terminologically complex. Indeed, the subtlety of Ainsworth's argument may have contributed to her decision not to attempt to publish the article, despite its influence on her students and collaborators. Her proposal was that individual adaptation (long-term thriving) may result from adaptation (changing oneself in order to respond) to the environment. However, there are forms of adaptation (long-term thriving) where refusal to adapt (change oneself in order to respond) is optimal—for instance, in depleting or punitive environments that can be changed or exited. Some forms of adaptation (thriving) may come at the expense of other forms of adaptation (thriving), as in the familiar case in which the demands of one area of life—family, work—come at the expense of others—diet, exercise, self-care. A further complexity lies in the fact that there are forms of adaptation (responding and/or thriving) that are based very directly on adaptation (species-level natural selection), such as the deployment of conditional strategies as evolutionary-based behavioural repertoires. However, there are forms of adaptation (responding and/or thriving) that are more based on social learning or other processes based more on human plasticity than responses directly grounded in adaptation (species-level natural selection).

Working closely with Mary Ainsworth was her former student Mary Main. Ainsworth and Main had observed that in the Strange Situation procedure a sizeable minority of infants directed attention away from their caregiver on reunion. These were infants whose primary caregiver had more frequently rebuffed their attempts to seek closeness when distressed (see Chapter 3). Main argued that, from an evolutionary perspective, avoidance could be interpreted as a proactive response by the infant that 'paradoxically permits whatever proximity is possible under conditions of maternal rejection'.²⁵ Evolutionary processes would have selected for avoidance as one part of the infant repertoire for responding to caregivers, because infants who are able to avoid antagonizing their caregivers or making demands that their caregiver will rebuff are more likely to have survived. An infant successfully utilizing an avoidant strategy maintains an indirect but real proximity to their caregiver, as well as the regulatory control to continue to be responsive to the environment.²⁶ Main argued that avoidance represented a behavioural repertoire, selected by evolutionary processes, and available to infants to respond to less sensitive care. Following use of the term in studies of animal behaviour, Main termed avoidance a "conditional strategy", since it would be drawn on when the primary attachment strategy of proximity-seeking is unavailable to an infant, and it would provide a conditional form of access to proximity.

Main anticipated that there would be two possible conditional strategies. One of these, avoidance, directed attention away from cues that might activate attachment behaviour. The other, appearing as ambivalent/resistant attachment in the Strange Situation, entailed vigilant attention to signs of the caregiver's potential unavailability. Both conditional strategies could be predicted, on average in human evolutionary history, to have provided sufficient indirect proximity with caregiver to have facilitated survival. In her 1984 paper, reflecting on Main's theory, Ainsworth stated that it may be that avoidance hinders adaptation (long-term thriving), or could be beneficial even in the long-run if conditions continue to be difficult.

²⁵ Main, M. (1979). 'The "Ultimate" Causation of Some Infant Attachment Phenomena: Further Answers, Further Phenomena, Further Questions'. *Behavioral and Brain Sciences*, 2: 640–643, p. 643.

²⁶ Main, M. (1981). 'Avoidance in the Service of Proximity: A Working Paper', in K. Immelmann, B. Barlow, L. Petrovich, and M. Main (eds), *Behavioral Development: The Bielefeld Interdisciplinary Project*, Cambridge: Cambridge University Press, pp. 694–699, p. 686.

Ainsworth offered her conviction that this was essentially an empirical matter: the question of whether 'avoidant attachment may be adaptive according to ultimate criteria in the mental health sense is clearly a researchable proposition'.²⁷

Main's stance was that, even if short-term benefits could be identified, avoidance would hinder long-term thriving. Yet a greater threat was represented by the potential for breakdown or disruption of attachment strategies, disorganized attachment, which were argued to be elicited by a child's experiences of frightened, frightening, or dissociative caregiving. Another student of Ainsworth's, Patricia Crittenden, held a different position. She argued that there may be ecological niches where conditional strategies are simply superior, especially under conditions—most of human evolutionary history—where significant danger is prevalent. To characterize conditional strategies as a second-best option then would be both overgeneralized and potentially ethnocentric.²⁸ Crittenden agreed with Main that there are two general forms of conditional strategy. She argued that Ainsworth's avoidant attachment classification represented a local case of the broader strategy of inhibiting or distorting information about negative emotions. Ainsworth's ambivalent/resistant classification represented a local case of the broader strategy of inhibiting or distorting information about the temporal and causal sequencing of others' availability, permitting the maintenance of vigilance. Crittenden proposed that with maturation, and responding to experiences of threat or danger, additional forms of these two strategies become available.²⁹ To the extent that symptoms are prompted by close relationships and/or experiences of threat, she suggested that many forms of mental illness may be regarded as effects of conditional strategies. For instance, one kind or component of post-traumatic stress disorder (PTSD) is avoidant symptoms, numbness, and other inhibitions of negative emotions. Another kind or component of PTSD is hyperarousal and vigilance for threats or potential separations from attachment figures.³⁰

In the 1990s, Fonagy was impressed by Main's account of conditional strategies as behavioural repertoires made available by human evolutionary history. This model aligned with Anna Freud's suggestion that many forms of mental illness begin as a local response to adverse circumstances, and that they remain relevant and perhaps even helpful so long as those adversities continue. However, Main's account added to Freud's an underpinning ethological–evolutionary framework, in which certain profiles of symptoms have an underpinning logic and connection in their likelihood of contributing to survival in the face of adversity. Yet, in recent years, Fonagy has increasingly advocated for Crittenden's position, which he has described as 'more inspiring'.³¹ Crittenden's account of the conditional

²⁷ Ainsworth, M. (1984). 'Attachment, Adaptation and Continuity'. Paper presented at the International Conference on Infant Studies, April, John Bowlby Archive, Wellcome Collections, PP/Bow/J.1/57.

²⁸ Crittenden, P. M. (2000). 'A Dynamic-Maturational Model of the Function, Development, and Organization of Human Relationships', in R. S. L. Mills, and S. Duck (eds), *Developmental Psychology of Personal Relationships*, New York: Wiley, pp. 199–218.

²⁹ Crittenden, P. M. (1997). 'Truth, Error, Omission, Distortion, and Deception: The Application of Attachment Theory to the Assessment and Treatment of Psychological Disorder', in S. M. C. Dollinger and L. F. DiLalla (eds), *Assessment and Intervention across the Lifespan*, Hillsdale, NJ: Erlbaum, pp. 35–76.

³⁰ Crittenden, P. M. (1997). 'Toward an Integrative Theory of Trauma: A Dynamic-Maturational Approach', in D. Cicchetti and S. Toth (eds), *The Rochester Symposium on Developmental Psychopathology*, Volume 10, Risk, Trauma, and Mental Processes, Rochester, NY: University of Rochester Press, pp. 34–84; Crittenden, P. M. and Heller, M. B. (2017). 'The Roots of Chronic PTSD: Childhood Trauma, Information Processing, and Self-Protective Strategies', *Chronic Stress*, 1: 1–13.

³¹ Fonagy, P. (2013). 'Commentary on "Letters from Ainsworth: Contesting the 'Organization' of Attachment"'. *Journal of the Canadian Academy of Child & Adolescent Psychiatry*, 22(2): 178–179: 'Most would agree that, clinically, Crittenden's approach is more inspiring, particularly in understanding the behaviour of children and young people whose life has been blighted by malevolence on the part of their carers' (p. 179).

strategies aligned with the Fonagy and Luyten characterization of affective and cognitive ‘poles’ of mentalization (see Chapter 4), and with Luyten’s distinction between avoidant and anxious forms of depression.³² Fonagy appreciated Crittenden’s claim that, when adversities continued, a conditional strategy may have distinct advantages for an individual, as well as significant costs. This aligned with Anna Freud’s approach to thinking about defences, and appeared to Fonagy to be a destigmatizing perspective.³³ Fonagy was also interested by the idea that even symptoms of PTSD such as hypervigilance may, in certain ways, serve as coping strategies when they become woven into day-to-day living, part of how an individual works out their characteristic response to challenges.³⁴

Another attachment theorist, Jay Belsky, had likewise argued that conditional strategies may have advantages within certain ecological niches. Belsky speculated that humans evolved to treat the early care we receive from attachment figures as a ‘signal’ about the safety or danger of the environment. He suggested that sensitive care prompts a secure attachment relationship, reflecting confidence in others and their availability in the expectation of a more favourable environment. Insensitive care prompts an insecure attachment relationship, reflecting caution about trust in others and their availability in the expectation of a less favourable environment. An individual’s endocrinology, behaviour, and forms of information processing may then be calibrated by a concern with survival and short-term benefits, even if these come with a long-term price. Following developmental adversities, on average, individuals may be anticipated to be more impulsive, anxious, distractible, and more ready to engage in sexual and aggressive behaviours. Seeking to examine this hypothesis, Belsky conducted secondary analyses on the data from large cohort studies. He found that attachment insecurity was associated with earlier date of menarche, and harsh early care associated with more sexual risk-taking behaviours.³⁵ Belsky interpreted these findings as suggesting that the conditional strategies may have conferred benefits under adversity in human evolutionary history. They prompt early puberty and earlier sexual behaviour, increasing the likelihood in evolutionary history that a woman would survive to reproduce, though this strategy is also costly. For instance, early menarche is associated with long-run disadvantages

³² Luyten, P. and Blatt, S. J. (2012). ‘Psychodynamic Treatment of Depression.’ *Psychiatric Clinics*, 35(1): 111–129. See also Rost, F., Luyten, P., and Fonagy, P. (2018). ‘The Anaclitic–Introjective Depression Assessment: Development and Preliminary Validity of an Observer-Rated Measure.’ *Clinical Psychology & Psychotherapy*, 25(2): 195–209.

³³ Fonagy, P. (1999). ‘Points of Contact and Divergence between Psychoanalytic and Attachment Theories.’ *Psychoanalytic Inquiry*, 19(4): 448–480: ‘Crittenden’s (1990) work has been particularly helpful in translating behaviors typical of avoidant and resistant attachment patterns into the language of the defensive behaviors of infancy’ (p. 452); Fonagy, P. (2016). ‘The Role of Attachment, Epistemic Trust and Resilience in Personality Disorder: A Trans-Theoretical Reformulation.’ DMM News, 26 September. Accessed at: <http://www.iasa-dmm.org/images/uploads/DMM%20%2322%20Sept%2016%20English.pdf>: ‘The model’s emphasis on attachment as a means of protecting the self and one’s offspring from danger is then used to make a valuable—and in the process, valuably destigmatising—account of dysfunction as arising from knowledge acquired through environmental influence’ (p. 2).

³⁴ See e.g. Fonagy, P. (2012). ‘The Neuroscience of Related Trauma and Evidence-Based Intervention.’ Accessed at: <https://smad6740.files.wordpress.com/2015/01/the-neuroscience-of-related-trauma-and-evidence-based-intervention1.pdf>: Such ‘anomalies could confer a short-term advantage: Vigilance to threat; It is found in healthy soldiers exposed to combat. But they constitute a latent neural risk that predisposes to an increased likelihood of maladaptation in safe contexts (e.g. school) and of adult psychopathology.’

³⁵ Belsky, J., Houts, R. M., and Fearon, R. P. (2010). ‘Infant Attachment Security and the Timing of Puberty: Testing an Evolutionary Hypothesis.’ *Psychological Science*, 21(9): 1195–1201; Belsky, J., Steinberg, L., Houts, R., and Halpern-Felsher, B. (2010). ‘The Development of Reproductive Strategy in Females.’ *Developmental Psychology*, 46(1): 120–128; Belsky, J. (2019). ‘Early-Life Adversity Accelerates Child and Adolescent Development.’ *Current Directions in Psychological Science*, 28(3): 241–246. See also Szepeswol, O. and Simpson, J. A. (2019). ‘Attachment within Life History Theory: An Evolutionary Perspective on Individual Differences in Attachment.’ *Current Opinion in Psychology*, 25: 65–70.

in terms of physical health. Early sexual debut may be linked to less discrimination of sexual partners and less safety in the relationship.

While Crittenden and Belsky frame their claims as criticism of Main's stance, Fonagy felt that they reflected different levels of analysis: "To my mind, in the same way that light can be seen as either waves or particles, the consequences of attachment trauma can be seen as an adaptation that also reflects the absence of an organised strategy. I see no loss of meaning coming from this admittedly heuristic or rather deeper integration of these models."³⁶ Main's primary concern was with motivational systems, their modulation or disruption; Belsky's was with factors that might have helped past humans survive to the age of reproduction and see offspring survive; Crittenden's was with the potential benefits of the inhibition or distortion of emotion or cognition under conditions of threat. Fonagy agreed with Main that childhood experiences of maltreatment or trauma can result in breakdowns or interruptions of emotion regulation and attentional control. These are important contributors to mentalization, so problems in these areas are likely to have significant ramifications for an individual in their experience of self and future social interactions. However, Fonagy agreed with Belsky that 'natural selection may have solved the problem of environmental variability by using the attachment system as an "early warning system" to indicate to an infant the degree to which violent conduct may be required later in life.'³⁷ A whole variety of social, genetic, and endocrinal calibrations maybe made as a result of the early warning provided by childhood, priming later responses.³⁸ Finally, Fonagy agreed with Crittenden that disruptions of emotion regulation and attentional control may be patterned in ways that have specific advantages under conditions of adversity.³⁹

Even emotions like fear can contribute to the development of habitual ways of responding that benefit an individual in certain ways.⁴⁰ PTSD symptoms of hyperarousal offer a good illustration. Fonagy and colleagues held that Main was right that these symptoms can be disruptive for an individual's capacity to regulate their feelings and maintain attention and responsiveness to their immediate environment. These symptoms may also be psychologically and physiologically very costly. Yet the potential for hyperarousal may have contributed to human survival in our evolutionary history, where dangers may have been frequent and pressing. If the present likewise contains significant potential dangers then, Fonagy agreed with Crittenden that hyperarousal can have localized advantages for an individual in cuing a rapid response to potential threats. If the present does not contain relevant dangers, then

³⁶ Fonagy, P. (2013). 'Commentary on "Letters from Ainsworth: Contesting the 'Organization' of Attachment"'. *Journal of the Canadian Academy of Child & Adolescent Psychiatry*, 22(2): 178–179, p. 179.

³⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 62.

³⁸ Fonagy, P. and Campbell, C. (2017). 'What Touch can Communicate: Commentary on "Mentalizing Homeostasis: The Social Origins of Interoceptive Inference" by Fotopoulou and Tsakiris', *Neuropsychoanalysis*, 19(1): 39–42: 'The attachment relationship therefore serves as an indicator of the nature of the infant's environment (Belsky, Steinberg, & Draper, 1991; Simpson & Belsky, 2016), indeed it appears to be a powerful communication mechanism which works at the level of gene expression as well as at the level of social cognition, as Meaney's work (Meaney, 2010; Meaney & Szyf, 2005) and some epigenetic human studies have shown' (p. 40).

³⁹ E.g. Fonagy, P. and Sharp, C. (2008). 'Treatment Outcome of Childhood Disorders: The Perspective of Social Cognition', in C. Sharp, P. Fonagy, and I. Goodyer (eds), *Social Cognition and Developmental Psychopathology*, New York: Oxford University Press, pp. 411–470: 'All behavioural-emotional problems of children are both a reflection of biological vulnerability and/or adversity and an effort to cope and adapt' (p. 448).

⁴⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 62: 'The basic emotions [curiosity, fear, anger, lust, caring, sadness and joy] are survival and adaptive experiences, which are essential and inform action and reaction. Many patients will try to avoid them but in doing so will diminish their ability to appraise situations and themselves' (p. 307).

indeed, he agreed with Main, hyperarousal is likely overwhelmingly counterproductive. In making appeal to the concept of adaptation, Fonagy has been accused by critics of absorbing the problems of earlier ego psychology in identifying mental health with subjection to hegemonic social norms.⁴¹ There is an extent to which this criticism holds, but it does not have full purchase. Fonagy's use of the concept of adaptation has been qualified by acknowledgement, especially in recent years, regarding the harms of social norms that institutionalize non-mentalizing and prompt adaptations from individuals that hinder their long-term creativity and thriving (Chapter 9).

From the perspective of Fonagy's approach to theorizing adaptation, even mental health symptoms that reduce environmental responsiveness, like dissociation, may be appraised for the extent to which they developed in response to past adversity, and the extent to which they are reinforced by helping an individual respond to their current challenges.⁴² Fonagy has argued that both appraisals are important in order to understand what is maintaining mental health problems. Clinicians who intervene without this understanding may risk increasing their client's suffering and reducing their capacity to cope with their life. Fonagy likewise argued that the forms of non-mentalizing may similarly have evolved as repertoires because they contributed to survival under conditions of adversity. According to this logic, where circumstances would specifically penalize mentalizing, it may risk harm to patients to encourage its development. However, Fonagy and colleagues also argue that, when the environment is safe enough to mentalize adversity and trauma, this can prove grist to the mill. Skills gained in processing these experiences can lead to the specific development of strengths in reconsidering the thoughts and feelings of oneself and others, prompting the phenomenon of post-traumatic growth.⁴³ For instance, Fonagy and colleagues have documented the potential for elevated self-reported empathy in adults who have experienced childhood trauma.⁴⁴ However, they caution that this elevated empathy is not yet well understood, and could reflect hypermentalization rather than mentalization of others, or both may be making a contribution.

In their 2017 paper, 'What We have Changed our Minds About' (see Chapter 4), Fonagy, Luyten, Allison, and Campbell applied these ideas about adaptation in thinking about borderline personality disorder (BPD). They argued that many mental disorders, BPD among them, are 'nested in the context of the evolutionary priorities of the human condition'.⁴⁵

⁴¹ Ferraro, D. ([2011] 2014). 'The Other, Clinical and Empirical: A Review of Fonagy et al. On Affect Regulation, Mentalisation, and the Development of the Self'. Accessed at: <https://melbournelacanian.wordpress.com/2014/11/24/the-other-clinical-and-empirical-a-review-of-fonagy-et-al-on-affect-regulation-mentalisation-and-the-development-of-the-self/>.

⁴² Drinkwater, M. (2008). 'Portraits of a Divided Self'. *Guardian*, 11 September. Accessed at: <https://www.theguardian.com/society/2008/sep/11/mentalhealth>: 'Peter Fonagy, a psychology professor at University College London, said: "Consciousness is one of the remaining mysteries of neuroscience. We may be able to explain the features of [dissociative identity disorder] DID, but without understanding consciousness we cannot entirely explain how DID occurs." Far from being considered pathological, Fonagy suggested that DID can be seen as a sign of resilience; a strategy for coping with extremely traumatic events. "The mind is capable of separating and we put traumatising events into little 'boxes'. This can be useful as it can help people dissociate from these experiences", he said.' See also Fonagy, P. (2018). 'Preface', in Werner Bohleber (ed.), *Destructiveness, Intersubjectivity and Trauma: The Identity Crisis of Modern Psychoanalysis*, London: Routledge, pp. xi–xiv.

⁴³ Fonagy, P. and Bateman, A.W. (2016). 'Adversity, Attachment, and Mentalizing'. *Comprehensive Psychiatry*, 64: 59–66: 'The experience of trauma, and understanding and overcoming adversity, could also be considered to entail the enhancement of mentalizing in the process of overcoming trauma' (p. 60).

⁴⁴ Greenberg, D. M., Baron-Cohen, S., Rosenberg, N., Fonagy, P., and Rentfrow, P. J. (2018). 'Elevated Empathy in Adults Following Childhood Trauma'. *PLoS One*, 13(10): e0203886.

⁴⁵ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 1. Borderline Personality Disorder as a Limitation of Resilience'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 11.

Following Belsky, Fonagy and colleagues argued that caregiving experiences represent a form of social communication about the most effective way for a child to function. They speculated that:

In certain situations, for example, an emergency milieu characterized by high levels of interpersonal aggression, the heightened and immediate sensitivity and seemingly instinctive and physically charged form of appraisal characteristic of BPD might in fact be adaptive, at least in the short term. In such an environment, extreme vigilance is a potential advantage, and similarly, the ability to form intense emotional relationships quickly might elicit resources or protection ... we postulate that this mentalizing profile may be a response to cues suggestive of an unreliable and potentially threatening social environment. We thus should be wary of seeing apparent dysfunctions of the clinically 'hard to reach' as indicative of a deficit or any kind of sub-optimal functioning (as, indeed, we have done previously). We would now consider that what may appear to us as dysfunction is an evolutionarily primed adaptation.⁴⁶

BPD is, here, regarded in a way that integrates the conclusions of Main, Belsky, and Crittenden. In part, it reflects breakdown of stabilizing and organizing aspects of emotion and attentional regulation, often but by no means always, following from adverse experiences of care in childhood and/or childhood trauma. This then contributes to non-mentalizing social interactions with others, which helps sustain difficulties in modulating emotion and attention, giving the impression of a disorder of 'personality'. In part, BPD reflects a repertoire pre-primed by human evolution, and elaborated on the basis of experience, in which early adversity and trauma function as signals that hyperarousal and a focus on short-term concerns will facilitate survival. The slow work of generating second-order representations of mental states and reappraising them is impatiently downplayed. And, in part, BPD reflects a profile of symptoms that may be reinforcing for an individual if it helps them cope with present adversities.⁴⁷

For example, in environments characterized by high levels of violence, it may be expected that security, trust, and the mentalization of oneself and others would result in risk of exploitation and increased danger—for instance, through a failure to access aggression when needed.⁴⁸ Functioning in teleological mode, by contrast, could be expected to achieve a more immediate response, with greater effectiveness at least in the short term. Bateman and Fonagy have claimed that it may be beneficial to 'prioritise an external, nonreflective, rapidly responding focus on the control of others'.⁴⁹ In fact, it should be acknowledged that longitudinal

⁴⁶ Ibid.

⁴⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'particularly in these individuals, subsequent adversity or trauma could further disrupt mentalising, in part as an adaptive manoeuvre on the part of the individual to limit exposure to a brutalising psychosocial environment, and in part because the high levels of arousal generated by attachment hyperactivation and disorganised attachment strategies serve to disrupt less well-practiced and less robustly established higher cognitive capabilities' (p. 47).

⁴⁸ Bevington, D., Fuggle, P., Cracknell, L., and Fonagy, P. (2017). *Adaptive Mentalisation-Based Integrative Treatment: A Guide for Teams to Develop Systems of Care*, Oxford: Oxford University Press: 'In the face of very real social conflict, then a mentalising stance in relation to social relationships is likely only to leave the individual more vulnerable; for example, it may ultimately render the individual less able to make use of aggressive and violent social strategies, which might seem absolutely necessary for self-protection. In such an environment, more reactive, impulsive forms of mentalising and non-mentalising modes of behaviour, which may most conspicuously take the form of aggression or risky sexual behaviour, might be at one level an adaptive response' (p. 63).

⁴⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 440.

evidence to date has generally shown the opposite: that under conditions of ongoing adversity, secure attachment in childhood is an asset rather than a deficit for later thriving.⁵⁰ However, it may be that studies have not been designed to identify the specific benefits that might stem from non-mentalizing.⁵¹ Or that the studies have not been conducted in contexts with the levels of adversity and violence Fonagy and colleagues think are relevant.

Epistemic trust

Drawing on these reflections about adaptation, Fonagy and colleagues have argued that ‘as clinicians, we may end up calling these individuals “hard to reach,” yet they are simply showing what may be a reasonable adaptation to a social environment where information from most attachment figures is “tagged” as likely to be misleading.’ As such, ‘it may be more accurate to characterize BPD as an “emergency mode” form of social understanding.’⁵² The defining characteristic of this ‘emergency form’ of social understanding is a higher bar for trusting the claims and experiences of others. Trust has been a topic of rapidly rising interest in the social sciences since the 1990s, in a social context in which management of vulnerability and complexity has become a central concern.⁵³ The question of trust in professionals—including psychological professionals—has reflected, in intensified form, this wider concern, because precisely part of what it is to be a professional is to provide a service that depends in part on trust.⁵⁴ Fonagy and colleagues were also reflecting on empirical findings that showed that early abuse and neglect predicted later symptoms of personality disorder, with mediation by reflective function clearly in operation, but to less of an extent than anticipated.⁵⁵ Other factors besides mentalization appeared to be important for the relationship between early adversity and personality disorder.

Drawing from the work of Sperber and colleagues,⁵⁶ Fonagy and colleagues have termed ‘epistemic vigilance’ an attitude in which the claims and experiences of others are not felt to be dependable, generalizable, or relevant.⁵⁷ They conceptualized epistemic vigilance as the

⁵⁰ E.g. Sroufe, L. A., Egeland, B., Carlson, E. A., and Collins, W. A. (2005). *The Development of the Person*. New York: Guilford Press; Englund, M. M., Kuo, S. I. C., Puig, J., and Collins, W. A. (2011). ‘Early Roots of Adult Competence: The Significance of Close Relationships from Infancy to Early Adulthood.’ *International Journal of Behavioral Development*, 35(6): 490–496. Short-run benefits of insecure attachment under adversity, in this case maternal depression, have been documented by Milan, S., Snow, S., and Belay, S. (2009). ‘Depressive Symptoms in Mothers and Children.’ *Developmental Psychology*, 45(4): 1019–1033. However, the researchers found that there were longer-term advantages to security under conditions of adversity.

⁵¹ Ein-Dor, T., Mikulincer, M., Doron, G., and Shaver, P. R. (2010). ‘The Attachment Paradox: How can so Many of Us (the Insecure Ones) Have no Adaptive Advantages?’. *Perspectives on Psychological Science*, 5(2): 123–141.

⁵² Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). ‘What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication.’ *Borderline Personality Disorder and Emotion Dysregulation*, 4: 9.

⁵³ Giddens, A. (1991). *Modernity and Self-Identity*, Cambridge: Polity Press; Sievers, B. (2003). ‘Against all Reason: Trusting in Trust.’ *Organizational & Social Dynamics* 3(1): 19–39.

⁵⁴ Gorman, E. H. and Sandefur, R. L. (2011). “Golden Age,” Quiescence, and Revival: How the Sociology of Professions Became the Study of Knowledge-Based Work. *Work & Occupations*, 38(3): 275–302. Already in 2000, Fonagy was arguing that ‘Therapy is based on trust, and it is hard for it to be effective without it.’ Interview with Fonagy, cited in Rice, M. (2000). ‘Therapy is the New Religion.’ *Guardian*, 13 August. Accessed at: <https://www.theguardian.com/theobserver/2000/aug/13/life1.lifemagazine11>.

⁵⁵ Chiesa, M. and Fonagy, P. (2014). ‘Reflective Function as a Mediator between Childhood Adversity, Personality Disorder and Symptom Distress.’ *Personality and Mental Health*, 8(1): 52–66.

⁵⁶ Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., and Wilson, D. (2010). ‘Epistemic Vigilance.’ *Mind & Language*, 25(4): 359–393.

⁵⁷ A different formulation is offered by Asen, E. and Fonagy, P. (2020). ‘Mentalization in Systemic Therapy and Its Empirical Evidence’, in Matthias Ochs, Maria Borcsa, and Jochen Schweitzer (eds), *Systemic Research in*

default state of humans, a valuable tool to protect against misinformation stemming from malice or incompetence on the part of others: 'All young humans are at the mercy of a knowledge differential, uncertain about the trustworthiness of the information they are about to receive ... We are more likely to survive and thrive if we do not forget that not everyone will teach us things that are true or indeed in our interest to learn.'⁵⁸ It should be noted that there are three important differences between Sperber's original claims and the exposition of them offered by Fonagy and colleagues, which was mediated by Gergely's interpretation:⁵⁹

- 1) Whereas for Sperber, epistemic vigilance is a 'suite of cognitive mechanisms',⁶⁰ the characterization of epistemic vigilance as an 'attitude' by Fonagy and colleagues conveys the impression of a more unitary phenomenon.
- 2) Whereas for Sperber, epistemic vigilance seeks to assign the right amount of credence to people and conversations, this meaning remains operative, but Fonagy and colleagues also give the term the sense of distrust. For instance, Fonagy and Allison proposed that disorganized attachment means that a 'child seeks others to confirm or deny his/her own understanding, which he/she has little faith in, but, being unable to trust information received from others, remains in a state of uncertainty and epistemic vigilance.'⁶¹
- 3) Sperber insists that epistemic vigilance may characterize individuals, but can also characterize groups, institutions, and cultures. Yet even individual epistemic vigilance occurs on the basis of repertoires that may be suggested or imposed by the individual's wider community. Whether at an individual or a collective level, epistemic vigilance is then achieved through what Hutchins calls 'distributed cognition', a network of processes that together create effects greater than the sum of their parts.⁶² For instance, Sperber identifies that epistemic vigilance may be sustained by the coordination of 'cognitive artefacts (e.g. measuring instruments), techniques (e.g. statistical tests of confidence),

Individual, Couple, and Family Therapy and Counseling, New York: Springer, pp. 207–221. There epistemic trust is conceptualized as treating information as possessing 'authenticity and personal relevance' (p. 216), rather than as 'dependable, generalisable or relevant'. This emphasis on authenticity is interesting, and materially distinct from the idea of dependable and generalizable information. Yet another definition is offered in Fisher, S., Guralnik, T., Fonagy, P., and Zilcha-Mano, S. (2020). 'Let's Face It: Video Conferencing Psychotherapy Requires the Extensive Use of Ostensive Cues'. *Counselling Psychology Quarterly*, Early View: 'Epistemic vigilance refers to the ability to be justifiably suspicious about socially transmitted information, in a manner that helps protect one against potentially deceitful or erroneous information.'

⁵⁸ Fonagy, P. and Allison, E. (2018). 'The Origin of Human Life: A Psychoanalytic Developmental Perspective', European Psychoanalytical Federation, 31st Annual Conference, Warsaw, 24 March.

⁵⁹ Gergely, G. (2008). 'Learning "about" versus Learning "from" Other Minds: Natural Pedagogy and its Implications', in P. Carruthers, S. Laurence, and S. Stich (eds), *The Innate Mind. Foundations and the Future*, Volume 3, New York: Oxford University Press, pp. 170–198. Speber became a colleague of Gergely at the Department of Cognitive Science at the Central European University in 2010, following retirement from CNRS, Paris. Gergely's interpretation of Sperber and colleagues' concept of epistemic vigilance pre-dates this, however.

⁶⁰ Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origi, G., and Wilson, D. (2010). 'Epistemic Vigilance'. *Mind & Language*, 25(4): 359–393, p. 359. See also Sperber, D. and Wilson, D. (2015). 'Beyond Speaker's Meaning'. *Croatian Journal of Philosophy*, 15(44), 117–149.

⁶¹ Fonagy, P. and Allison, E. (2014). 'The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship'. *Psychotherapy*, 51(3): 372–380, p. 374.

⁶² Hutchins, E. (1995). *Cognition in the Wild*, Cambridge, MA: MIT Press; Hutchins, E. (2006). 'The Distributed Cognition Perspective on Human Interaction', in Nicholas Enfield and Stephen C. Levinson (eds), *Roots of Human Sociality: Culture, Cognition and Human Interaction*, Oxford: Berg, pp. 375–398; Hutchins, E. (2010). 'Cognitive Ecology'. *Topics in Cognitive Science*, 2: 705–715. For a distributed cognition perspective on playful exploration, for example, see Parker-Rees, R. (2014). 'Playfulness and the Co-Construction of Identity in the First Years', in L. Brooker, M. Blaise, and S. Edwards, (eds), *The SAGE Handbook of Play and Learning in Early Childhood*, London: SAGE, pp. 366–377.

and procedures (e.g. for cross-examination).⁶³ By contrast, Fonagy and colleagues have tended, except in their most recent work (see Chapter 9), to limit epistemic vigilance to individuals. They have also neither acknowledged nor explored the distributed cognition that Sperber and colleagues believe underpins epistemic vigilance, with the exception of some initial remarks on the affordances of mobile phone technology (see Chapter 9).

Comparing the work of Sperber and colleagues with the use of it by Fonagy and colleagues, the differences are quite marked. Most importantly, the concept of epistemic vigilance as used by Fonagy and colleagues slides between two very different meanings, without this distinction being marked. Sometimes, Fonagy and colleagues use the term to mean ‘discrimination’, the opposite of blind trust.⁶⁴ This occurs especially when epistemic vigilance is distinguished from ‘epistemic hypervigilance’ and ‘epistemic mistrust’. Perhaps more commonly, however, the term ‘epistemic vigilance’ is taken to mean, precisely, hypervigilance and mistrust. For Fonagy and colleagues, ‘epistemic vigilance’ often appears to mean a state in which information from others is not felt to have bearing or resonance. This reduces exposure to others’ potential malice or incompetence, at the price of knowledge that would indeed be accurate and useful. In epistemic vigilance then, the betrayal and disappointment we dread are already within us to an extent, in the way that our dread of being undermined by others ultimately undermines our own capacity to benefit from encounters with others.⁶⁵

In the clinic, an inability to relax epistemic vigilance can be seen when, ‘no matter how true or accurate the therapist’s interpretations are, the patient will not be able to make use of them because they are not experienced as true’, at least in any genuine and personally relevant sense.⁶⁶ The patient cannot learn from experience—or modify their thoughts, feelings, or behaviour on the basis of this learning—because they are unable to make use of potential inflows of information. Fonagy and colleagues therefore argued that ‘personality disorder is not a “disorder of personality” but an inaccessibility to cultural communication relevant to the self from the social context’.⁶⁷

Conversely, epistemic trust has been defined by Fonagy and colleagues as ‘an individual’s willingness to consider new knowledge from another person as trustworthy, generalizable, and relevant to the self’;⁶⁸ or as ‘openness to acquiring social knowledge that is regarded as personally relevant and of generalizable significance’.⁶⁹ As such, epistemic trust delineates those whose claims and experiences we (feel we can) rely upon in order to understand the

⁶³ Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., and Wilson, D. (2010). ‘Epistemic Vigilance’. *Mind & Language*, 25(4): 359–393, p. 383.

⁶⁴ The term ‘discrimination’ is from Bion, W. R. (1976). ‘Interview with Anthony G. Banet’, in *Tavistock Seminars*, London: Karnac Books, pp. 97–114: ‘To what extent is one to allow any idea to come in? One feels that there’s a need for a sort of discriminating screen’ (p. 113). A related distinction between ‘default’ and blind’ trust is made in Granqvist, P. (2020). *Attachment in Religion and Spirituality: A Wider View*, New York: Guilford Press.

⁶⁵ In significant regards, the thinking of Fonagy and colleagues about epistemic mistrust echoes earlier thinking about the origins of prejudice in the breakdown of trust within attachment relationships as a prompt for ‘suspicion’ in internal working models. See Fonagy, P. and Higgitt, A. (2007). ‘The Development of Prejudice: An Attachment Theory Hypothesis Explaining its Ubiquity’, in H. Parens, A. Mahfouz, S. W. Twemlow, and D. Scharff (eds), *The Future of Prejudice: Psychoanalysis and the Prevention of Prejudice*, Lanham, MD: Rowman & Littlefield, pp. 63–79.

⁶⁶ Allison, E. and Fonagy, P. (2016). ‘When is Truth Relevant?’. *Psychoanalytic Quarterly*, 85(2): 275–303, p. 285.

⁶⁷ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). ‘What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication’. *Borderline Personality Disorder and Emotion Dysregulation*, 4: 9, p. 27.

⁶⁸ Fonagy, P. and Allison, E. (2014). ‘The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship’. *Psychotherapy*, 51(3): 372–380, p. 373.

⁶⁹ Fisher, S., Guralnik, T., Fonagy, P., and Zilcha-Mano, S. (2020). ‘Let’s Face It: Video Conferencing Psychotherapy Requires the Extensive Use of Ostensive Cues’. *Counselling Psychology Quarterly*, Early View.

world and to change ourselves to effectively respond to it.⁷⁰ According to Fonagy and Allison, writing in 2014, ‘epistemic trust is there to ensure that the individual can safely change their position; it triggers the opening of what we can think of as an “epistemic superhighway”—an evolutionarily protected mechanism that signals readiness to acquire knowledge.’⁷¹ It is not stated whether, like mentalization, this epistemic trust depends on imaginative capacities, as a kind of suspension of disbelief (see Chapter 5). Presumably it does, unless the knowledge transmitted is wholly procedural rather than subjective.⁷² As Winnicott argued, and Fonagy and colleagues have echoed in other contexts, cultural knowledge needs to be both subjectively conceived as well as objectively discovered if it is to be alive for a person, and usable.⁷³

In articulating the developmental basis for epistemic trust, Fonagy and colleagues criticized psychoanalytic theories that sought to account for a child’s experience of the minds of others in terms of ‘projective identification’; the term is too overlaid with different meanings to convey precisely what is taking place.⁷⁴ They also criticized those, such as Trevathan, who account for a child’s experience of the minds of others in terms of ‘intersubjectivity’.⁷⁵ Again an over-encompassing term, intersubjectivity implies that mental states are ‘shared’. Fonagy and colleagues have generally regarded mental states as only ever the properties of individuals, though in recent writing they have brought this assumption into question (see Chapter 9). Rather than appeal to projective identification or ideas of intersubjectivity, Fonagy and colleagues instead drew on the work of Gergely and Csibra to propose that epistemic trust tends to be activated by interactions in which we feel the other to recognize our i) mental states, ii) intentionality, and iii) individuality. The concrete behaviours through which this sense of recognition is conveyed were termed ‘ostensive cues’ by Gergely and Csibra.⁷⁶ For Fonagy and colleagues, ‘ostensive cues indicate to the infant that the caregiver

⁷⁰ There remains some instability in the account of epistemic trust as to whether it refers to felt experience or a causal mechanism or both. For Fonagy’s emphasis on the importance of this distinction for psychoanalytic theory, see Chapter 6.

⁷¹ Fonagy, P. and Allison, E. (2014). ‘The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship’. *Psychotherapy*, 51(3): 372–380, p. 374.

⁷² The extent to which imagination is implicated in procedural information has been debated by philosophers. Ultimately, it hinges on the definition of ‘imagination’ and the model of perceptual experience. Matherne, S. (2016). ‘Kant’s Theory of the Imagination’, in A. Kind (ed.), *Routledge Handbook of Philosophy of Imagination*, London: Routledge, pp. 55–68.

⁷³ Winnicott, D. W. (1971). *Playing and Reality*, London: Routledge. See the discussion of Winnicott in Fonagy, P. and Target, M. (1996). ‘Playing with Reality: I. Theory of Mind and the Normal Development of Psychic Reality’. *The International Journal of Psychoanalysis*, 77: 217–233.

⁷⁴ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 20.

⁷⁵ Fonagy, P., Gergely, G., and Target, M. (2007). ‘The Parent–Infant Dyad and the Construction of the Subjective Self’. *Journal of Child Psychology and Psychiatry*, 48(3–4): 288–328. See also Gergely, G. and Csibra, G. (2006). ‘Sylvia’s Recipe: The Role of Imitation and Pedagogy in the Transmission of Cultural Knowledge’, in Nicholas Enfield and Stephen C. Levinson (eds), *Roots of Human Sociality: Culture, Cognition and Human Interaction*, Oxford: Berg, pp. 229–255; Gergely, G. and Unoka, Z. (2008). ‘Attachment, Affect-Regulation, and Mentalization: The Developmental Origins of the Representational Self’, in C. Sharp, P. Fonagy, and I. M. Goodyer (eds), *Social Cognition and Developmental Psychopathology*, New York: Oxford University Press, pp. 305–342.

⁷⁶ Csibra, G. and Gergely, G. (2011). ‘Natural Pedagogy as Evolutionary Adaptation’. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567): 1149–1157; Egyed, K., Király, I., and Gergely, G. (2013). ‘Communicating Shared Knowledge in Infancy’. *Psychological Science*, 24(7): 1348–1353. Subsequent research has suggested that Csibra and Gergely overstated the necessary role of ostensive cues for joint attention and for learning: e.g. Gredebäck, G., Astor, K., and Fawcett, C. (2018). ‘Gaze Following is not Dependent on Ostensive Cues: A Critical Test of Natural Pedagogy’. *Child Development*, 89(6): 2091–2098, though the argument that ostensive cues may facilitate engagement with the mind of the other is unaffected by these qualifications. A further point to raise is that Csibra and Gergely ostensibly take the term ‘ostensive cues’ from their reading of Sperber and Wilson. However, comparison of the two groups of theorists suggests that the former have reified the notion, because for the latter it is only a designation for a set of a broader class of strategies through which attentional processes are directed within the dance of conversation. It would also appear that, for Sperber and Wilson, ostensive

recognises him or her as an individual, and as a thinking and feeling (i.e. mentalizing) “agent”.⁷⁷ Fonagy sometimes also refers to ostensive cues as offering acknowledgement of the other’s experience of ‘personal agency’, which is given no definition, but presumably signifies *the felt state in which mental states, intentionality, and individuality have been acknowledged*. Ostensive cues are characterized as diverse and may include joint attention, turn taking, social referencing, and marked mirroring.⁷⁸ Kindness has been conceptualized as an ostensive cue.⁷⁹ One form of kindness, gentle touch that acknowledges and responds to embodied cognition, is an especially powerful ostensive cue.⁸⁰ Fonagy and colleagues argued that humans have evolved to treat receipt of ostensive cues as a basis for epistemic trust, since because they signal the presence of another mind interested in our intentions, thoughts, and/or feelings, and therefore likely reflect a situation of relative safety and social supports.⁸¹

In the developmental model proposed, parental reflective function leads to the display of ostensive cues towards a child. In turn, Fonagy and colleagues claimed that this process predisposes a secure attachment relationship, because recognition of the intentions and/or mental state of the child will facilitate secure base and safe haven provision. At times, Fonagy and colleagues appear to imply that parental reflective function leads to a quantitative increase in ostensive cues towards the child. However, this does not appear to be their intended position. Over-use or ill-judged use of ostensive cues such as touch and eye contact would rather be intrusive caregiving, which has been associated with insecure and disorganized attachment rather than secure attachment.⁸² Instead, it must be assumed that ostensive cues are used in a ‘sensitive’ way: responding accurately to the signals and communications of the child in a way that is experienced by the child as appropriate and prompt. This would imply a reconciliation of the opposition drawn by Fonagy and colleagues between reflective function

cues do not necessarily imply recognition of the other’s intentional agency: it depends on what is meant by the latter concept, which can be rather opaque (see Chapter 9).

⁷⁷ Bateman, A. and Fonagy, P. (2019). ‘Introduction’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 3–20, p. 16.

⁷⁸ Fonagy and colleagues have recently published interesting reflections on how ostensive cues may be achieved through video working platforms such as Zoom—for instance, using joint attention through the screen share function. Fisher, S., Guralnik, T., Fonagy, P., and Zilcha-Mano, S. (2020). ‘Let’s Face It: Video Conferencing Psychotherapy Requires the Extensive Use of Ostensive Cues’. *Counselling Psychology Quarterly*, Early View; Fonagy, P., Campbell, C., Truscott, A., and Fuggle, P. (2020). ‘Mentalising Remotely: The AFNCCF’s Adaptations to the Coronavirus Crisis’. *Child and Adolescent Mental Health*, 25(3): 178–179. Attempts to deliver therapy remotely to young people in their homes also suggested to the researchers that another ostensive cue professionals may use is behaviour indicating that a message was for the recipient specifically rather than others in earshot.

⁷⁹ Fonagy, P. (2020). ‘Kindness can Work Wonders. Especially for the Vulnerable’. *Guardian*, 17 May. Accessed at: <https://www.theguardian.com/society/2020/may/17/kindness-can-work-wonders-especially-for-the-vulnerable>.

⁸⁰ Fonagy, P. and Campbell, C. (2017). ‘What Touch can Communicate: Commentary on “Mentalizing Homeostasis: The Social Origins of Interoceptive Inference” by Fotopoulou and Tsakiris’, *Neuropsychoanalysis*, 19(1): 39–42.

⁸¹ Anna Freud has been given as an example by Allison and Campbell: ‘in accounts of her interactions with children, one gets the strong sense that Anna Freud was tremendously gifted at eliciting epistemic trust in the children she worked with, and that she achieved this through her genuine recognition of and interest in the minds of the children she worked with’. Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre. See descriptions in Kennedy, H. (1983). ‘Anna Freud 1895–1982’. *Psychoanalytic Quarterly*, 52(4): 501–506; Solnit, A. J. and Newman, L. M. (1984). ‘Anna Freud: The Child Expert’, *The Psychoanalytic Study of the Child*, 39: 45–63; Blomfield, O. H. (1991). ‘Anna Freud: Creativity, Compassion, Discipline’. *International Review of Psycho-Analysis*, 18: 37–52.

⁸² Carlson, E. A. (1998). ‘A Prospective Longitudinal Study of Attachment Disorganization/Disorientation’. *Child Development*, 69(4): 1107–1128; Wang, F., Cox, M. J., Mills-Koonce, R., and Snyder, P. (2015). ‘Parental behaviors and beliefs, child temperament, and attachment disorganization’. *Family Relations*, 64(2): 191–204.

and sensitivity discussed in Chapter 3, and account for their large shared variance in accounting for child attachment security.

Yet, at the same time as contributing to attachment security, Fonagy and colleagues argue that (appropriate use of) ostensive cues facilitate development of a capacity to lower epistemic vigilance under appropriate circumstances.⁸³ Both secure attachment and the capacity for epistemic trust are, Fonagy and colleagues suggest, resources for social adaptation and mentalization. They suppose that the attachment system evolved to ensure the protection of juveniles by having them seek their familiar caregiver when alarmed. However, the attachment system may, in human evolutionary history, have also gained the additional function of calibrating the capacity for epistemic trust, to equip the individual for the relative safety or danger of the environment in which they will grow up.⁸⁴ In relative safety, the capacity for epistemic trust is a great social asset. In relative danger, its risks increase.

These reflections raise the question of the exact relationship between attachment and epistemic trust, not least because one way that Ainsworth had characterized individual differences in the Strange Situation was precisely in terms of the infant's trust in their caregiver.⁸⁵ In recent years, Ainsworth's remarks have been developed further by attachment researchers; there has been a growing trend for secure attachment to be re-described as a kind of trust in the caregiver's availability.⁸⁶ In 2014, Fonagy and Allison attempted a re-description of all the Strange Situation classifications in these terms. Secure attachment was conceptualized as well-calibrated trust in the other. Ambivalent/resistant attachment was interpreted as excessive credulity to the perspective of the other; avoidant attachment was interpreted as vigilance about the perspective of others. Disorganized attachment was characterized as a disruption in the foundational basis of trust in others, prompting a state of persistent epistemic vigilance.⁸⁷ Fonagy and Allison claimed support for this stance from a study by Corriveau, Harris, Meins, and colleagues. The researchers gave children stimuli to categorize, with input from their caregiver and a stranger. Children from securely attached dyads relied

⁸³ Fonagy, P. and Allison, E. (2014). 'The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship'. *Psychotherapy*, 51(3): 372–380: 'While attachment may be a key mechanism for mediating epistemic trust, it is secondary to an underlying biological process preserved by evolution. In other words, secure attachment is unlikely to be necessary for generating epistemic trust but it may be sufficient to do so, and, further, it is the most pervasive mechanism in early childhood because it is a highly evolutionarily effective indicator of trustworthiness' (p. 375).

⁸⁴ For discussion of this claim, see Granqvist, P. (2020). 'Attachment, Culture, and Gene-Culture Co-Evolution: Expanding the Evolutionary Toolbox of Attachment Theory'. *Attachment & Human Development*, 23(1): 90–113. Alongside praise for the advance represented by attention to epistemic trust in evolutionary perspective, Granqvist ultimately criticizes Fonagy and colleagues for neglecting the contribution of epistemic trust to cultural transmission and evolution. On theory of mind as an exaptation of the phylogenetically earlier attachment system, see also Chisholm, J. S. (2003). 'Uncertainty, Contingency, and Attachment: A Life History Theory of Theory of Mind', in K. Sterelny and J. Fitness (eds), *From Mating to Mentality: Evaluating Evolutionary Psychology*, New York: Psychology Press, pp. 125–153.

⁸⁵ Ainsworth characterized ambivalent/resistant patterns as representing distrust in the caregiver. Ainsworth, M. D. S., Bell, S. M., and Stayton, D. J. (1974). 'Infant-Mother Attachment and Social Development: "Socialisation" as a Product of Reciprocal Responsiveness to Signals', in J. M. Richards (ed.), *The Integration of a Child into a Social World*, Cambridge: Cambridge University Press, pp. 9–135: 'It may be viewed as advantageous for an infant whose mother seems to him to move unpredictably and inconsistently (and whom he has not been able to learn to trust) to monitor her movements with exceptional alertness and to evince disturbance whenever she moves off' (p. 125).

⁸⁶ E.g. Verhees, M. W., Ceulemans, E., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., De Winter, S., and Bosmans, G. (2017). 'The Effects of Cognitive Bias Modification Training and Oxytocin Administration on Trust in Maternal Support: Study Protocol for a Randomized Controlled Trial'. *Trials*, 18(1): 326: 'Bowlby proposed that whether children are able to develop trust in parental support depends for a significant part on children's experiences of care in response to distress during interactions with their parents. Through repeated interactions with sensitive and responsive parenting, children develop trust in the parent and become securely attached' (p. 2).

⁸⁷ Fonagy, P. and Allison, E. (2014). 'The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship'. *Psychotherapy*, 51(3): 372–380, p. 374.

flexibly on the advice of their caregiver and the stranger, depending on who was giving the most plausible information. Children from avoidantly attached dyads tended to ignore the input from their caregiver even when the advice was relevant and helpful. Children from anxious/ambivalent dyads preferred information from their caregiver when stimuli were 50/50 ambiguous. Children from dyads that received a disorganized attachment classification showed a preference for the mother's advice when the stimulus was novel.⁸⁸ The data, however, do not line up neatly with the position of Fonagy and colleagues. The account of disorganized attachment as disruption of trust and the generation of epistemic vigilance finds little support in the study by Corriveau and colleagues. If anything, the findings weakly point in the opposite direction. Furthermore, elsewhere Fonagy and colleagues have associated both ambivalent/resistant (or preoccupied) and disorganized (or unresolved) attachment with epistemic vigilance and BPD. This is poorly aligned with the finding by Corriveau and colleagues that ambivalent/resistant attachment was associated with apparent trust in the mother when stimuli could otherwise just as easily be classified one way as the other.

Another relatively unconvincing model of the relationship between attachment and epistemic trust was presented by Fonagy and colleagues in their 'What We have Changed our Minds About' paper (see Figure 7.1).

In this model, only the history of adversities feeds mentalizing problems, with difficulties in emotion regulation and attachment contributing only through this adversity. This does not seem plausible: the concepts of 'adversity' and 'dysfunction' are losing purchase, as they become overinvested. In the model presented in 'What We have Changed our Minds About', mentalizing problems are described as causing persistent epistemic vigilance through two routes. A first is through the failure to pick up on and use others' ostensive cues as a prompt for epistemic trust in them. A second is through general loss of interest in social communication. In this account, there is no contribution of epistemic mistrust to difficulties in mentalizing; it is assumed that causality is only in the other direction.

A more recent, and more persuasive, characterization of the relationship between epistemic trust and attachment was presented by Fonagy and colleagues at the Epistemic Petrification Conference at Bristol University in 2018. In a keynote address, Fonagy and colleagues updated the 1998 and 2008 models (see Chapter 3) of the role of caregiving in the development of mentalization (see Figure 7.2).

First, Fonagy and colleagues described a pathway in which the capacity for epistemic trust was supported. Sensitive caregiving was defined by Ainsworth as the caregiver's ability to 'perceive and to interpret accurately the signals and communications implicit in her infant's behavior, and given this understanding, to respond to them appropriately and promptly'.⁸⁹ Fonagy and colleagues argued that sensitive care is in large part underpinned by parental reflective function, which allows the caregiver to give ostensive cues to the child through behaviours that perceive and respond to signals and communications suggesting the child's mental states. This facilitates the child's development of secure attachment and epistemic trust. In the 2018 keynote, Fonagy and colleagues also permitted a reinforcing pathway from secure attachment to epistemic trust. Use of the caregiver as a secure base may facilitate exploration of things and people. But to find learning in the social aspects of exploration,

⁸⁸ Corriveau, K. H., Harris, P. L., Meins, E., Fernyhough, C., Arnott, B., Elliott, L., ... and De Rosnay, M. (2009). 'Young Children's Trust in their Mother's Claims: Longitudinal Links with Attachment Security in Infancy'. *Child Development*, 80(3): 750–761.

⁸⁹ Ainsworth, M. (1969). 'Sensitivity vs. Insensitivity to the Baby's Signals Scale'. Accessed at: http://www.psychology.sunysb.edu/attachment/measures/content/ainsworth_scales.html.

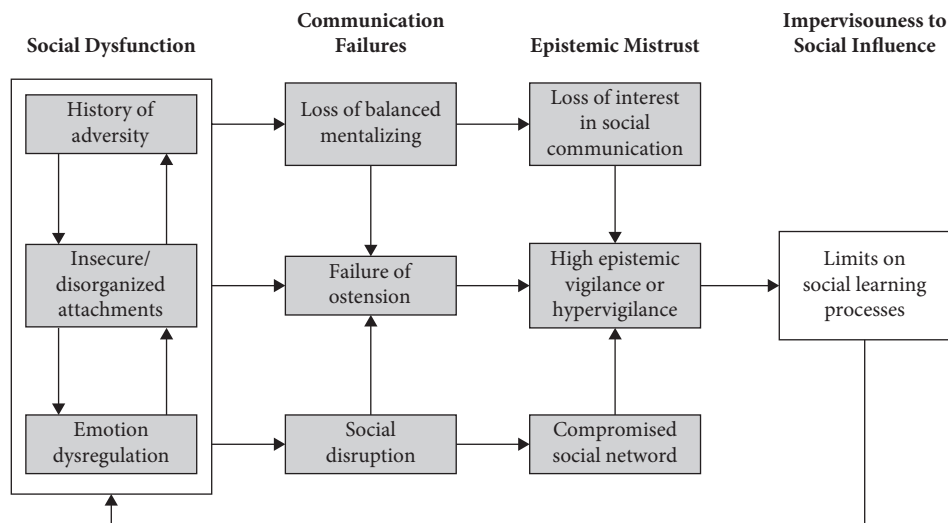


Figure 7.1 The Natural Pedagogy Model of Personality Disorder. Illustrates the interaction between social dysfunction, failure in social communication, epistemic mistrust, and imperviousness to social influence that underpin personality disorder. Emotion dysregulation, disrupted attachment histories and the disorganized insecure attachment system generate social/interpersonal dysfunction. This undermines accurate social communication, causing social disruption, the misinterpretation of social signals, and difficulty in recognizing ostensive cues from others. These difficulties in the area of social communication can give rise to epistemic mistrust in relation to the social environment. This is not inherently a maladaptive process: epistemic vigilance has a natural function. However, the absence of epistemic trust sets a limit upon social learning. This can render the individual potentially unable to function effectively within their social environment and can lead to further disruption in the social network, leaving the individual increasingly isolated and prone to further social/interpersonal dysfunction

Source: Reproduced from Peter Fonagy, Patrick Luyten, Elizabeth Allison, and Chloe Campbell, 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication', *Borderline Personality Disorder and Emotion Dysregulation*, 4(9): Figure 1, DOI: <https://doi.org/10.1186/s40479-017-0062-8> Copyright © The Authors. Licensed under a Creative Commons Attribution 4.0 International License.

and for this to be experienced as having genuine bearing and importance, epistemic trust is needed. As such, in the updated 2018 model, learning about the (social) world and the self are routed through epistemic trust.

The relationship between learning about the self, learning about the world, and mentalizing in the 2018 model is curious. Fonagy and colleagues retain from Fonagy's 1998 model the Hegelian position that learning about the self is mediated through learning about the world. But the 2018 model describes only partial mediation. This implies that an individual can engage in learning about the self through epistemic trust without learning about the world. This once more raises the unresolved question in Fonagy's thinking about whether envisaging and conceptualizing the i) mental states and ii) behaviour of the self and the iii) mental states and iv) behaviour of the other can feasibly be characterized under the same umbrella, and whether one or another element has priority. In Fonagy and Luyten's 2009 paper, the Hegelian account was still dominant, with experience of the mental states of

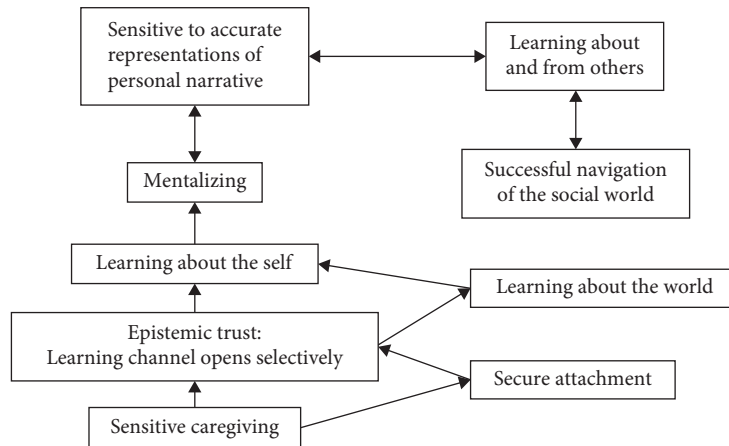


Figure 7.2 The role of caregiving in the development of mentalization.

Source: Reproduced from Fonagy, P. Allison, Luyten, P. and Campbell, C. (2018). 'Epistemic Petrification'. Keynote address at the Epistemic Petrification Conference, Bristol University, 2–3 July.

others given developmental priority, and experience of the mental states of oneself occurring on the basis provided by the former. In the 2018 model, this mediation clearly remains important. But there seems to be a direct route to learning about the self as well. Furthermore, whereas the 1998 diagram suggested that a mentalized account of the caregiver's behaviour contributes to the development of mentalizing capacity, in the 2018 diagram there is no line between learning about the world and mentalizing. The implication would be that learning about the self has some priority, then, for the initial development of mentalizing in early childhood.

However, as we saw in Chapter 6, the term 'self' is a complex one in the vocabulary of Fonagy and his collaborators. It may be suspected that the meaning of 'self' in the phrase 'learning about the self' is the individual's experience as an embodied subject, since other aspects of 'self', such as the sense of personal agency, have clearly been theorized elsewhere as depending on mentalization (see Chapter 6). This is supported by the fact that the production of a self-representation that weaves together our experiences, wishes and plans, self-knowledge, and other facets of the 'self' is distinguished from 'learning about the self' and placed in a bidirectional relationship with mentalizing. Hegel returns here: learning about and from others is placed in a bidirectional relationship with mentalizing after early childhood, mediated by an individual's narratively constructed sense of self and self-relevance. It is this learning that supports successful navigation of the social world, and mentalizing is important to the extent that it facilitates such learning, not in itself.

In the 2018 keynote at the Epistemic Petrification Conference, Fonagy and colleagues also presented a further diagram (Figure 7.3), representing a developmental model from problems with early care to persistent use of non-mentalizing modes.

Here, neglect or traumatic attachment relationships contribute to the maintenance of epistemic vigilance and to disorganized attachment.⁹⁰ In the 2018 keynote, Fonagy and colleagues stated that they were not sure whether disorganized attachment contributed to the maintenance

⁹⁰ Chiesa and Fonagy had found that abuse and neglect made independent contributions to personality disorder. Chiesa, M. and Fonagy, P. (2014). 'Reflective Function as a Mediator between Childhood Adversity, Personality Disorder and Symptom Distress'. *Personality and Mental Health*, 8(1): 52–66.

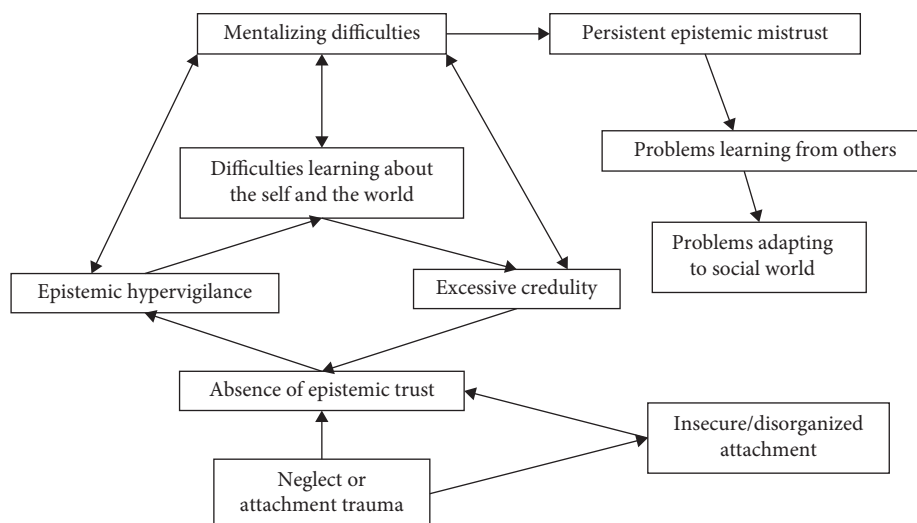


Figure 7.3 Developmental model from problems with early care to persistent use of non-mentalizing modes.

Source: Reproduced from Fonagy, P. Allison, Luyten, P. and Campbell, C. (2018). 'Epistemic Petrification'. Keynote address at the Epistemic Petrification Conference, Bristol University, 2–3 July.

of epistemic vigilance. This was in contrast to an earlier version of the same diagram, presented in various places in 2017 including London and San Francisco, in which the pathway from disorganized attachment to epistemic trust was absent.⁹¹ Difficulties sustaining epistemic trust contribute to difficulties in social learning. Unlike in the model of the development of mentalizing, in which learning about the self and the world were distinguished as playing somewhat different roles, in the development of non-mentalizing their role is characterized as much the same. In turn, this can contribute to excessive credulity at times, because inaccuracies in the perceptions of others' behaviour and one's own mind can lead the misleading claims of others to be mistaken for accurate ostensive cues about one's mental states.⁹² There are also bidirectional links placed between epistemic hypervigilance and problems with mentalizing: an important difference from the model in the 'What We have Changed our Minds About' paper, in which there was no path at all from epistemic vigilance to difficulties in mentalizing.⁹³

⁹¹ Fonagy, P. (2017). 'The Concept of Epistemic Trust: Can it Help Integrate Attachment Theory?'. Paper presented to the San Francisco Centre for Psychoanalysis, Haskell Norman Prize Event, Monday, 9 January 2017. The diagrams also featured in the keynote address, 'Epistemic Trust and Attachment', to the International Attachment Conference, London, 1 July 2017. There seems to be some differences among Fonagy's collaborators on how attachment theory is conceived and the extent to which early attachment is regarded as critical for later development. For instance, we see claims such as Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Mentalising and Trauma', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 79–102: 'the development of the stress system ... is largely determined by the quality of attachment relationships' (p. 82).

⁹² This argument appeared already in Fonagy, P. and Target, M. (2000). 'Playing with Reality: III. The Persistence of Dual Psychic Reality in Borderline Patients'. *The International Journal of Psychoanalysis*, 81(5): 853–873: 'The patient lacks an authentic, organic self-image built around internalised representations of self-states. The absence or weakness of such a self-image leaves the child, and later the adult, with affect that remains unlabelled and confusing, uncontained (Bion, 1962) ... This will create a desperation for meaning, and a willingness to take in reflections from the other that do not map on to anything within the child's own experience' (p. 865).

⁹³ A certain hesitancy regarding the directionality of relations between mentalizing and epistemic trust can also be seen in Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Assessment of Mentalising', in Anthony

The 2018 keynote by Fonagy and colleagues also proposed bidirectional relationships between credulity and problems with mentalizing. Non-mentalizing prompts feelings of isolation and loneliness, because the individual's mind is not able to take nourishment from their interactions with others. This stimulates 'epistemic hunger', in which the strength of the desire to feel recognized by others may reinforce credulity when someone seems to offer the possibility of recognition.⁹⁴ Elsewhere during the Epistemic Petrification Conference, Fonagy clarified how this credulity would occur:

If I do not have any clear view of myself then when anyone says something about me, I may experience this as a match, and excessive trust can develop ... So somebody says 'you're a bad person' and I have some image of myself as evil. I may trust that statement; that's very common for the patients that I treat. An analogy comes in with extremist politics here: you have a popular set of ideologies that are clearly defensive. They say 'I'm being maltreated. The world is unjust to me. I have been excluded.' Then you have someone who wants to acquire their trust, who says 'yes, you have been excluded. You have been maltreated. You should be rebelling and I'm surprised that you've put up with this for so long.' A populist ideology, and they come to be trusted. That is a sinister pattern seen in politics today in many places.⁹⁵

In the first case, Fonagy describes the experience of match between social denigration and feelings of self-denigration. This is a situation he regards as especially predisposed today by the circulation of idealized images of lifestyle and success in social media, which can offer a picture of the self as bad in comparison. On the one hand, difficulties learning about the self and forms of non-mentalizing may facilitate this experience of match. As may adolescence, as a developmental stage in which cognitive and social supports for achieving clarity and specificity in self-representation are not fully online. On the other hand, the experience of social denigration matching feelings of self-denigration can be anticipated to contribute to further difficulties in mentalizing about oneself—for instance, by reinforcing both distress and deployment of psychic equivalence in the form of certainty about the self's worthlessness. In

Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 37–62: 'The assessor should determine to what extent mentalising impairments are linked to problems with epistemic trust' (p. 50).

⁹⁴ Fonagy, P., Allison, E., and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 72. Fonagy, P. and Allison, E. (2018). 'The Origin of Human Life: A Psychoanalytic Developmental Perspective', European Psychoanalytical Federation, 31st Annual Conference, Warsaw, 24 March: 'Epistemic hypervigilance and epistemic credulity, can co-exist and can oscillate rapidly. This can be confusing for the clinician whose understanding appears to be valued and idealized one day only to be scotomised, spat out, denigrated and spurned the next.'

⁹⁵ Duschinsky, R., Collver, J., and Carel, H. (2019). 'Trust Comes from a Sense of Feeling One's Self Understood by Another Mind': An Interview with Peter Fonagy', *Psychoanalytic Psychology*, 36(3), 224–227, p. 225. See also Luyten, P., Campbell, C., and Fonagy, P. (2019). 'Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach', *Journal of Personality*, 88(1): 88–105: 'We understand such excessive epistemic credulity as being triggered by a hyperactive or unmoored imagination generating a personal narrative that is too diffuse to provide an accurate sense of differential awareness of others' capacity to perceive oneself. Excessive credulity results, as all personal narratives feel as if they "fit" sufficiently for trust to be generated, making the person vulnerable to exploitation.' The most detailed case study of epistemic credulity offered by Fonagy and colleagues, though without using the term, is the case of Mr A. in Fonagy, P. and Target, M. (2005). 'Some Reflections on the Therapeutic Action of Psychoanalytic Therapy', in J. Auerbach, K. Levy, and C. E. Shaffer (eds), *Relatedness, Self-Definition and Mental Representation: Essays in Honor of Sidney J. Blatt*, New York: Taylor & Francis, pp. 191–212.

the second case, Fonagy describes the experience of political solicitation. This is, in Fonagy's view, an extreme and nasty use of a skill characteristic of all effective influencers: the 'massive difference in ability of individuals to influence (teachers, politicians, managers)' may be regarded as 'explicable in terms of varying capacity to generate epistemic trust'.⁹⁶ In the political case under discussion, the politician's solicitation creates the illusion of recognizing the listener as an intentional subject and their feelings of having been wronged. This may be facilitated by forms of non-mentalizing, such as psychic equivalence, in which feelings are conflated with political reality. In turn, the match can be anticipated to contribute to forms of non-mentalizing. This might include difficulties in experiencing apparent political opponents as having minds, or even being fully human like oneself; a requirement for urgent action in teleological mode; and/or aspects of pretend mode in which political ideology is treated as reality, without attention to its actual limitations in characterizing states of affairs.

In Fonagy's earlier work, both affect regulation and attentional control had been described as requisites and/or component elements and/or effects of mentalization. He had also elsewhere theorized bidirectional relationships between mentalization and social support—for instance, in helping an individual reframe and digest difficult experiences.⁹⁷ The 2018 diagram did not include affect regulation, attentional control, or social support. But, in presenting his keynote, Fonagy verbally situated that they contributed to mentalizing difficulties. The focus in the keynote was on another relevant component, which was difficulties learning about the self and the world. When these difficulties prompt non-mentalizing, this helps sustain a state of epistemic mistrust, blocking the capacity to experience the claims and experiences of others as trustworthy and relevant. In turn, problems in adapting to the social world are likely to ensue. This will depend in part on the extent to which epistemic mistrust becomes context-insensitive.⁹⁸

An interesting point for the model, raised by Fonagy and colleagues elsewhere, is that it is not just abuse and neglect that can lead to epistemic distrust, though they regard it as an especially important route. Social and political forms of adversity may predispose the development of epistemic distrust, as may various genetic and temperamental factors.⁹⁹ Fonagy and colleagues have also argued that epistemic distrust is not simply characteristic of pathology, but is a feature of all our lives at times. As we saw in Chapter 6, Fonagy and colleagues have argued that sexuality and aggression are intentions that caregivers will, to some extent, reject or ignore rather than acknowledge. Fonagy and Allison observed that 'this leaves us all with a gap, a failed definition, a rupture of epistemic trust'.¹⁰⁰ The cycle of epistemic hypervigilance and epistemic credulity can therefore be expected to occur with aspects of

⁹⁶ Fonagy, P. (2013). 'Mentalizing and the Attachment Process'. Paper delivered at the Menninger Clinic, 22 March. Accessed at: <http://www.creatingconnections.nl/assets/files/2013.04.18%20Peter%20Fonagy%20Creating%20Connections.pdf>.

⁹⁷ E.g. Luyten, P., Van Assche, L., Kadriu, F., Krans, J., Claes, L., and Fonagy, P. (2017). 'Other Disorders often Associated with Psychological Trauma', in C. Dalenberg, S. Gold, and J. Cook (eds), *APA Handbook of Trauma Psychology*. Volume 1: Foundations in Knowledge, Washington, DC: American Psychological Association, pp. 243–280, p. 246.

⁹⁸ A 'preliminary' distinction between state-like and trait-like problems with epistemic trust has been put forward in Fisher, S., Guralnik, T., Fonagy, P., and Zilcha-Mano, S. (2020). 'Let's Face It: Video Conferencing Psychotherapy Requires the Extensive Use of Ostensive Cues'. *Counselling Psychology Quarterly*, Early View.

⁹⁹ Luyten, P., Campbell, C., and Fonagy, P. (2019). 'Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach'. *Journal of Personality*, 88(1): 88–105.

¹⁰⁰ Fonagy, P. and Allison, E. (2015). 'A Scientific Theory of Homosexuality for Psychoanalysis', in A. Lemma and P. E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, London: Routledge, pp. 125–137, p. 134. See also Lorenzini, N., Campbell, C., and Fonagy, P. (2019). 'Mentalisation and its Role in Processing Trauma', in Bernd Huppertz (ed.), *Approaches to Psychic Trauma: Theory and Practice*, Lanham, MD: Rowman & Littlefield, pp. 403–422.

the alien self, and perhaps with particular frequency in relation to sexual desire and aggressive intentions. Given that these are among the major challenges of adolescence, it is perhaps then unsurprising that ‘an average adolescent is in a state of often quite acute epistemic mistrust’,¹⁰¹ though punctuated by epistemic credulity with peer groups and sometimes more generally.¹⁰² To the extent that epistemic hypervigilance and credulity come to dominate an individual’s experiences of sexuality or aggression, it can be anticipated that they will stop being able to take in and process information about themselves and the world.

Marinus van IJzendoorn and Marian Bakermans-Kranenburg, eminent attachment researchers, also presented a paper at the Epistemic Petrification Conference. They raised a variety of critical points following on from the 2018 keynote paper by Fonagy and colleagues.¹⁰³ A first was that, naturally for emergent scientific work, the model remains highly speculative. There is, as yet, no validated measure of epistemic trust.¹⁰⁴ Though there are many measures of trust and distrust, Fonagy and colleagues would need to decide whether they are articulating something more specific, and how it relates to existing constructs. Recently Bo, Bateman, and Kongerslev used the ‘trust in peers’ sub-scale of the self-report Inventory of Parent and Peer Attachment as a measure of epistemic trust. Yet, with items like ‘I trust my friends’, it is not clear that this addresses epistemic trust specifically.¹⁰⁵ No measure has been developed to assess epistemic vigilance. No measure has been developed to assess epistemic credulity. Some of the elements of the model are also quite vague, and will be difficult to operationalize without further articulation—for instance, the idea of ‘difficulties learning about self and the world’. Specific predictions about the relationship between caregiving and epistemic trust will need to be offered, likewise the relationship between epistemic trust and mentalization.

A second issue raised by van IJzendoorn and Bakermans-Kranenburg was that Fonagy and colleagues may be underestimating the multiple causal implications of attachment

¹⁰¹ Ezrati, O. (2014). ‘Freud Off: Giving New Meaning to Psychoanalysis’, Haaretz, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>.

¹⁰² See Fonagy, P., Target, M., Steele, M., Steele, H., Leigh, T., Levinson, A. et al. (1997). ‘Morality, Disruptive Behavior, Borderline Personality Disorder, Crime, and their Relationships to Security of Attachment’, in L. Atkinson and K. J. Zucker (eds), *Attachment and Psychopathology*, New York: Guilford Press, pp. 223–274, pp. 239–41. The extent to which adolescence is sufficiently different in this regard from adulthood as to require a different model of the relationship between mentalization and mental health has been raised by Battersby, S. (2018). *Is Mentalising Ability Associated with Mental Health Difficulties in Adolescents? A Systematic Review: Understanding the Construct of Mentalising in Adolescence and its Association with Mental Health, Structural Equation Model*. Unpublished doctorate in clinical psychology, University of Edinburgh.

¹⁰³ Van IJzendoorn, M. H. and Bakermans-Kranenburg, M. J. (2018). ‘Disorganized Attachment, Entropic Care, Epistemic Trust, and “Unusual” Beliefs’. Paper presented at the Epistemic Petrification Conference, 2–3 July, Bristol University.

¹⁰⁴ The Epistemic Trust Instrument was developed by Jennifer O’Connell in 2014. O’Connell, J. (2014). *Can We Develop an Adult Assessment Tool for Measuring Epistemic Trust?* Unpublished master’s thesis, London: University College London. However, this has yet to be used in any published study. Work is underway by the Heidelberg group to construct a measure of epistemic trust: Schröder-Pfeifer, P., Talia, A., Volkert, J., and Taubner, S. (2018). ‘Developing an Assessment of Epistemic Trust: A Research Protocol’. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 21(3).

¹⁰⁵ Bo, S., Bateman, A., and Kongerslev, M. T. (2019). ‘Mentalization-Based Group Therapy for Adolescents with Avoidant Personality Disorder: Adaptations and Findings from a Practice-Based Pilot Evaluation’. *Journal of Infant, Child, and Adolescent Psychotherapy*, 18(3): 249–262. The ‘trust in parents’ sub-scale of the self-report Inventory of Parent and Peer Attachment, taken at baseline, has also been found not to moderate reduction of symptoms in the course of therapy in a sample of adolescents. Such findings highlight the need for caution in assuming that existing measures of trust can be treated as measuring epistemic trust. Orme, W., Bowersox, L., Vanwoerden, S., Fonagy, P., and Sharp, C. (2019). ‘The Relation between Epistemic Trust and Borderline Pathology in an Adolescent Inpatient Sample’. *Borderline Personality Disorder and Emotion Dysregulation*, 6(1): 1–9. See also Gullone, E. and Robinson, K. (2005). ‘The Inventory of Parent and Peer Attachment—Revised (IPPA-R) for Children: A Psychometric Investigation’. *Clinical Psychology & Psychotherapy*, 12(1): 67–79.

relationships. Fonagy and colleagues were not sure whether or not disorganized attachment would contribute to epistemic dysfunction. However, this may be because they have assumed that disorganized attachment represents a single process (see Chapter 3). One process under the umbrella of disorganized attachment is dissociation and disorientation. Van IJzendoorn and Bakermans-Kranenburg asked what role epistemic hypervigilance and credulity might play in dissociative symptoms, and vice versa. As we saw in Chapter 4, in a 2017 paper Ensink, Fonagy, and colleagues criticized earlier work on mentalization for missing out the mediating role of dissociation. Van IJzendoorn and Bakermans-Kranenburg offered aligned concerns, but added that it would be valuable to assess dissociation and disorientation in developmental perspective and in light of the attachment relationship.

Van IJzendoorn and Bakermans-Kranenburg also wondered about the integrity of the construct of epistemic trust. They worried that the term ‘trust’ is ambiguous.¹⁰⁶ For instance, they observed that Fonagy and colleagues had discussed trust in the claims of others, and trust in their experience and viewpoint.¹⁰⁷ There could also be other objects of trust, such as trust in the other’s availability in an attachment relationship. Trust in others’ claims, trust in the usefulness or applicability of their experience and viewpoint, trust in their integrity, and trust in the other’s benevolence towards us may all vary relatively independently from one another.¹⁰⁸ Indeed, the ‘Epistemic Trustworthiness Inventory’ developed by Hendriks and colleagues distinguishes i) expertise/competence, ii) integrity, and iii) benevolence as separate scales.¹⁰⁹

To give an example of such differences in forms of trust: for an individual operating in teleological mode, a claim substantiated by hard evidence is persuasive; the other person’s experience and viewpoint are not (see Chapter 5). To offer another example: a child’s trust in their caregiver’s availability may well be somewhat helped by trust in the adult’s claims—for instance, about how best to make friends at school. But even if these claims turn out to have been mistaken, this need not have bearing for whether the child will feel that they can return to the caregiver as a safe haven when alarmed or distressed. For Van IJzendoorn and Bakermans-Kranenburg, it is the availability of the caregiver as a safe haven, rather than

¹⁰⁶ For discussions of the heterogeneous meanings of the concept of trust, see e.g. McKnight, D. H. and Chervany, N. L. (2001). ‘Trust and Distrust Definitions: One Bite at a Time’, in Rino Falcone, Munindar Singh, and Yao-Hua Tan (eds), *Trust in Cyber-Societies*, New York: Springer, pp. 27–54; Dietz, G. and Den Hartog, D. N. (2003). ‘Measuring Trust Inside Organisations’. *Personnel Review*, 32(5): 557–588; Joni, S.-N. (2004). ‘The Geography of Trust’. *Harvard Business Review*, 82(3): 7–11; Fink, M., Harms, R., and Möllering, G. (2010). ‘Introduction: A Strategy for Overcoming the Definitional Struggle’. *International Journal of Entrepreneurship and Innovation*, 11(2): 101–105; Driver, M. (2015). ‘How Trust Functions in the Context of Identity Work’. *Human Relations*, 68(6): 899–923; Duncan, G. (2019). *The Problem of Political Trust: A Conceptual Reformulation*, London: Routledge.

¹⁰⁷ In fact, Bateman and Fonagy appear to want to encompass all objects of trust. They described epistemic mistrust as entailing ‘destruction of trust in social knowledge of all kinds’. Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 27. Likewise, Luyten and colleagues give the example of a study in which ‘participants aged 12–17 who had experienced physical abuse were less able than their peers, who had no history of maltreatment, to correctly learn which stimuli were likely to result in reward, even after repeated feedback’. In Luyten’s characterization of the study, epistemic distrust is generic—it entails difficulty learning from any feedback on the properties of any stimuli. Luyten, P., Campbell, C., and Fonagy, P. (2019). ‘Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach’. *Journal of Personality*, 88(1): 88–105.

¹⁰⁸ The most important aspect of trust for Fonagy and colleagues appears to be less our conscious feelings about the other’s claims, and more whether we are willing to work this knowledge into our plans. This distinction is discussed by Holton, R. (1994). ‘Deciding to Trust, Coming to Believe’, *Australasian Journal of Philosophy*, 72: 63–76.

¹⁰⁹ Hendriks, F., Kienhues, D., and Bromme, R. (2015). ‘Measuring Laypeople’s Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI)’. *PLoS One*, 10(10): e0139309. Interestingly, the factor of ‘perceptions of the other as reliable/unreliable’ was excluded because it loaded on more than one dimension.

their provision of trustworthy communication about the social world, which can be anticipated to have the more significant role in facilitating affect regulation, attentional control, and future social competence, though trustworthy communication may no doubt contribute to the experience of a caregiver as a safe haven.

Similarly, mentalization may not always contribute to epistemic trust. Fonagy and colleagues specifically argued in 2007 that part of the reason mentalization evolved was to facilitate our capacity for deception: 'Mentalisation may serve competition: as the antlers of a reindeer are there to fight other reindeer rather than to catch prey or ward off predators. Self-awareness enables us to modify the way we wish to present ourselves, and to mislead.'¹¹⁰ However, this statement was from a period when they assumed mentalization was one thing, rather than a cluster of capacities. It may be that particular forms of mentalization are less facilitative of social deception (for instance, reconsideration of one's feelings in order to account for behaviours discrepant with our self-representation). The overarching concern for the 2018 model is that the relationship with experiences of care and with mentalization may be different, depending on the kind of trust and the kind of mentalization.

In the 2018 model of non-mentalizing, Fonagy and colleagues situated mentalizing difficulties as causing persistent epistemic mistrust, which in turn causes problems learning from others. But this claim could mean any number of things depending on what is meant by 'mentalizing difficulties', 'epistemic trust', and 'learning'. Consider, for instance, a situation in which mentalizing difficulties are taken to imply problems in conceiving of the mental states of others, epistemic mistrust is taken to imply problems in trusting the experiences and viewpoints of others, and learning from others is taken to mean gaining benefits from the thoughts of others. In this situation, we have done no more than state the logical implications of a problem in conceiving of the mental states of others. Consider, by contrast, a situation in which mentalizing difficulties are taken to mean problems in reconsidering one's own feelings as implicated in our behaviour, epistemic mistrust is taken to imply difficulty trusting the claims of others, and learning from others is taken to mean non-verbal imitation of valuable qualities in others. That mentalizing (in this sense) would cause epistemic mistrust (in this sense) would be a surprising, interesting, and falsifiable claim.

A further question left open by the 2018 keynote by Fonagy and colleagues is the relationship between 'absence of epistemic trust' and 'persistent epistemic mistrust'. Part of the problem, as has been widely acknowledged in the literature, is that the term 'trust' has various different meanings. Indeed, Lyon and colleagues have observed that 'trust is one of the most fascinating and fundamental social phenomena yet at the same time one of the most elusive and challenging concepts one could study.'¹¹¹ In Fonagy's 2018 diagram, '*absence of epistemic trust*' is prompted by neglect or attachment trauma (and perhaps insecure/disorganized attachment), whereas '*persistent epistemic mistrust*' is prompted by mentalizing difficulties and leads to problems learning from others.¹¹² They are not simply different names for the opposite of epistemic trust.

¹¹⁰ Fonagy, P., Gergely, G., and Target, M. (2007). 'The Parent-Infant Dyad and the Construction of the Subjective Self'. *Journal of Child Psychology and Psychiatry*, 48(3-4): 288-328, p. 297.

¹¹¹ Lyon, F., Möllering, G., and Saunders, M. N. K. (eds) (2012). *Handbook of Research Methods on Trust*. Cheltenham: Edward Elgar, p. 1.

¹¹² Only the epistemic vigilance that contributes to mentalizing difficulties, not the epistemic mistrust that results from mentalizing difficulties, figures in the diagram presented by Taubner, Gablonsky, and Fonagy in their account of the emergence of conduct disorder. This is despite the fact that it could readily be imagined that epistemic mistrust resulting from mentalizing difficulties could contribute to conduct problems. See Taubner, S., Gablonski, T.-C., and Fonagy, P. (2019). 'Conduct Disorder', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 301-321, Figure 18.1.

The 2018 diagram suggests that the relationship between absence of epistemic trust and persistent epistemic mistrust is mediated by mentalizing difficulties. Likewise, absence of epistemic trust, epistemic hypervigilance, and excessive credulity are all linked to problems learning from others only via the same route. There is no direct path from epistemic hypervigilance and/or excessive credulity to problems learning from others except via mentalization. This suggests differences between absence of epistemic trust and persistent epistemic mistrust in terms of their expected correlates. Though Fonagy and colleagues discuss epistemic trust as if it were a single construct, they seem to be describing two distinct processes—or at least two distinct facets of an umbrella process (see Table 7.1).

What Fonagy and colleagues term ‘absence of epistemic trust’ is described as contributing to mentalizing difficulties, whereas ‘persistent epistemic mistrust’ is fed by mentalizing difficulties. ‘Absence of epistemic trust’ is treated as causing difficulties learning about the self and the world; ‘persistent epistemic mistrust’ causes difficulties learning from others. ‘Absence of epistemic trust’ relates to epistemic vigilance in Sperber’s original sense: the concern to assign the right amount of credence to people and conversations is unable to be relaxed when needed. This leads to problems with mentalizing, by contributing to epistemic hypervigilance. ‘Persistent epistemic mistrust’ relates to epistemic vigilance in Fonagy’s revised sense: the experiences and claims of others are categorically distrusted. In both cases, the well-modulated and constructive doubt required for reconsideration of thoughts and feelings is hindered. For this reason, it is rather surprising that persistent epistemic mistrust is only ever an effect of problems with mentalizing, rather than also making a reciprocal contribution.

One explanation may be that, whereas the account of ‘absence of epistemic trust’ by Fonagy and colleagues seems to be primarily in dialogue with developmental psychology, his account of the ‘persistent epistemic mistrust’ is primarily in dialogue with personality and clinical psychology. The researchers’ characterization of ‘persistent epistemic mistrust’ suggests a state in which *the thoughts (and perhaps feelings) of others as implicated in their motivations and intentions are not felt to be trustworthy, generalizable, and relevant to the*

Table 7.1 Distinction between absence of epistemic trust and persistent epistemic mistrust

Psychological process	Absence of epistemic trust	Persistent epistemic mistrust
Apparent meaning	Inability to relax ordinary concern to assign the right amount of credence to people and conversations, blocking fast learning	The experiences and claims of others are categorically regarded as not to be trusted, generalized or regarded as relevant
Underpinned by	Epistemic vigilance (in Sperber and colleagues’ original sense, and one of the senses used by Fonagy and colleagues)	Epistemic vigilance (in Fonagy and colleagues’ distinct sense)
Key anticipated correlates	Causal contribution to i) non-mentalizing and to ii) difficulties in mentalizing	Caused by i) non-mentalizing, and ii) difficulties in mentalizing
Disciplinary context	Linguistics (specifically pragmatics), plugged into developmental psychology	Cognitive science (specifically natural pedagogy), plugged into personality and clinical psychology

self. Fonagy and colleagues situate mentalizing difficulties as mediating the effect of absence of epistemic trust on the capacity to learn from others. This may be because, in itself, extreme doubt or the absence of doubt about knowledge beyond perceptual experience do not represent a total block on learning from others, where this learning can figure within perceptual experience. What does present such a block is when others are not experienced as having thoughts (and feelings) distinct from our own perceptual experience, or when these are fabricated in pretend mode. Then there is nothing external from which to learn, even if others present claims about the credibility and relevance of their thoughts for us. Perceptual experience is experienced as immediate and immutable (psychic equivalence, teleological mode) or irrelevant (pretend mode), rather than as a contingent state that may be refined or modulated by the capacity to conceive or reconsider the thoughts and/or feelings of others. In general, mistrust is the expectation that others lack the competence or willingness to be reliable in acknowledging and meeting what we need or wish from them.¹¹³ Difficulties in conceiving or, especially, reconsidering the thoughts (and/or feelings) of others can be regarded as causing a specific form of mistrust, one in which there is little that can be learnt from the mental states of these others, since precisely these mental states are foreclosed or prejudged by non-mentalizing.

Fonagy and colleagues describe the effects of difficulties mentalizing others on the capacity to learn from others. Factors that have a predictable impact on mentalizing others, such as in-group/out-group dynamics, can be expected to likewise shape whether we are open to learning from these others. Missing, however, is attention to the effects of difficulties mentalizing the self on epistemic trust.¹¹⁴ Likewise, Fonagy and colleagues have as yet not been clear about whether epistemic trust/mistrust of feelings operates in the same way as epistemic trust/mistrust in thoughts. Asen, Campbell, and Fonagy have recently claimed that ‘critical thinking—some form of epistemic vigilance—is a highly valuable social cognitive resource’.¹¹⁵ This implies that the capacity to reconsider thoughts implicated in motivations and intentions of oneself can be facilitated by epistemic vigilance (this may link to the ‘unbinding’ function of meaninglessness discussed in Chapter 6). It is not clear whether there are forms of mentalization that are helped or hindered by epistemic distrust of others. However, Jurist has offered persuasive arguments that, with non-mentalizing and dehumanizing modes institutionalized and ascendant, epistemic distrust in contemporary culture and political discourse might facilitate mentalizing because it will facilitate reconsideration of the thoughts and feelings facilitated by this culture and discourse.¹¹⁶

Another outstanding question is what happens to mentalization when the self is subjected to persistent, as opposed to discriminated, epistemic mistrust. Luyten, Lemma, and Target have implied that this state characterizes depression, but this possibility has been raised only

¹¹³ Mollering, G. (2006). *Trust: Reason, Routine, Reflexivity*. New York: Elsevier; Barbalet, J. (2019). ‘The Experience of Trust: Its Content and Basis’, in Masamichi Sasaki (ed.), *Trust in Contemporary Society*, New York: Brill, pp. 11–30.

¹¹⁴ See e.g. Bion, W. ([1973] 1990). *Brazilian Lectures*, London: Karnac Books: ‘Some patients cannot . . . listen to what they themselves say. They have no respect for what they already know, so that their experience and knowledge are of no use to them. The question is not simply one of the relationship of the patient to the analyst, but the relationship of the patient with himself which may be so bad that he cannot even make use of what he already knows’ (p. 50).

¹¹⁵ Asen, E., Campbell, C., and Fonagy, P. (2019). ‘Social Systems: Beyond the Microcosm of the Individual and Family’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, p. 241.

¹¹⁶ Jurist, E. (2018). ‘Preface’, in *Minding Emotions: Cultivating Mentalisation in Psychotherapy*, New York: Guilford Press.

in passing.¹¹⁷ Another matter of remaining ambiguity is whether a capacity to effectively conceive of thoughts facilitates epistemic trust, epistemic vigilance, or either as needed in a discriminated way.¹¹⁸ Even with these questions still outstanding, the value ascribed by Asen, Campbell, and Fonagy to epistemic vigilance about thoughts as an asset to mentalization clearly disturbs any simple assumption that these are two ‘good things’ which must therefore come together. As we saw in Chapter 4, Bateman and Fonagy have described good mentalization of others in terms that seem to privilege considering their different thoughts and perspectives, whereas the qualities that characterize good mentalizing of oneself seem to privilege feelings and awareness of change in perspectives over time. This suggests that epistemic vigilance about the self’s present feelings would also be an asset to this form of mentalizing, in facilitating the reconsideration of inner sensations, affects, and moods, and the potential for feelings to be different across time. This vigilance may be especially what the Steeles and Fonagy were referring to in their initial concept of the ‘internal observer’, which predated the more general concept of ‘reflective function’ (see Chapter 3).

As discussed in Chapter 3, Fonagy and colleagues have headlined vaguely or variously defined terms like ‘mentalization’, which can help coalition-building among groups who can make various investments in the same buzzword, and thereby see themselves as holding the same stance and allegiance. The language of epistemic trust certainly appeals for this function: ‘trust’ is an undeniably potent term in contemporary society, and especially when addressed to clinical audiences for whom various kinds of trust will be salient and important.¹¹⁹ Trust is a language outlined in light, conveying shape more than contour and contrast. As such, the underspecification of epistemic trust may come at the price of conceptual confusion, especially regarding causal relationships.

To avoid this threat, firmer conceptual clarity is also needed on the distinction between epistemic vigilance, absence of epistemic trust, epistemic hypervigilance, and epistemic mistrust. Fonagy and colleagues have repeatedly emphasized that their thinking about epistemic trust is work in progress. In future work, it may be helpful to follow the usage by Fonagy and colleagues that maximally differentiates the concepts:

- ‘*Epistemic vigilance*’ would be returned to the usage of Sperber and colleagues, to mean a suite of cognitive mechanisms that help calibrate the credence assigned to the inferred experiences and claims of others. It might be added that such calibration is also relevant to our own inferred experiences and claims, given that these are rarely transparent to us in any simple way. Application of epistemic vigilance to our own thoughts and feelings entails generative doubt, and the potential for their reconsideration.
- When experience has led this suite of mechanisms to be applied too rigidly and conservatively, the result may be identified as ‘*absence of epistemic trust*’. This is no longer a suite of mechanisms but a quality: the inflexibility of those mechanisms. This quality may be more or less enduring. The corresponding attitude can be termed

¹¹⁷ Luyten, P., Lemma, A., and Target, M. (2019). ‘Depression’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 387–401, p. 398.

¹¹⁸ The latter option—mentalization facilitating discriminated trust/vigilance—seems to be implied, without being fully spelt out, in Fonagy, P., Campbell, C., and Allison, E. (2019). ‘Therapeutic Models Mentalising’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180.

¹¹⁹ Ellman, S. J. (2007). ‘Analytic Trust and Transference: Love, Healing Ruptures and Facilitating Repairs’. *Psychoanalytic Inquiry*, 27(3): 246–263.

‘epistemic hypervigilance’. This attitude has a bidirectional causal relationship with non-mentalizing.

- ‘*Epistemic mistrust*’ can be distinguished as a chronic state in which the thoughts (and perhaps feelings) of others as implicated in their motivations and intentions are not felt to be trustworthy, generalizable, and relevant to the self. This may appear as a fixed aspect of an individual’s ‘personality’, but in fact should be recognized as caused by non-mentalizing and the absence of mentalizing, which stabilize the state.

There is also further conceptual work to be done in specifying the relationship between forms of epistemic trust and forms of mentalizing, including in resolving problems that have been obscured by over-general uses of the two concepts.¹²⁰ Bateman, Fonagy, and Campbell have specified that teleological mode may prompt epistemic hypervigilance in a focus on moment-to-moment behaviour in lieu of trust in others’ stable intentions (see Chapter 5).¹²¹ And in earlier work Fonagy and Higgitt have described how a vicious cycle can arise, whereby psychic equivalence creates states of high arousal, which lead to reliance on pre-reflexive assumptions, which hinders the use of mentalizing, which sustains psychic equivalence.¹²² However, besides this, the relationship between epistemic hypervigilance and forms of non-mentalizing remains to be spelt out. Readers have tended to assume that Fonagy and colleagues believe that epistemic trust is an effect, not a cause, of mentalizing.¹²³ But the relationship is actually bidirectional and more complicated. The urgency of specification of this theory lies partly in the fact that over the coming years Fonagy and colleagues will be wishing to develop a measure of epistemic trust, to empirically test anticipated relationships. If this measure incorporates the present obscurity in the notion of epistemic

¹²⁰ One way of considering epistemic trust and mentalization is in terms of constituent elements and their inter-relation. Given the different definitions and uses Fonagy and colleagues have made of the two concepts, it would appear that there are eight permutations of epistemic trust to consider, against eight relevant permutations of mentalization:

Epistemic trust of the self’s feelings.
Epistemic trust of the self’s thoughts.
Epistemic trust of others’ feelings.
Epistemic trust of others’ thoughts.
Epistemic mistrust of the self’s feelings.
Epistemic mistrust of the self’s thoughts.
Epistemic mistrust of others’ feelings.
Epistemic mistrust of others’ thoughts.
Conceiving of the other’s feelings.
Conceiving of the other’s thoughts.
Conceiving of the self’s feelings.
Conceiving of the self’s thoughts.
Reconsidering the other’s feelings.
Reconsidering the other’s thoughts.
Reconsidering the self’s feelings.
Reconsidering the self’s thoughts.

¹²¹ Bateman, A., Fonagy, P. and Campbell, C. (2019). ‘Borderline Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 323–334, p. 328.

¹²² Fonagy, P. and Higgitt, A. (2007). ‘The Development of Prejudice: An Attachment Theory Hypothesis Explaining its Ubiquity’, in H. Parens, A. Mahfouz, S. W. Twemlow, and D. Scharff (eds), *The Future of Prejudice: Psychoanalysis and the Prevention of Prejudice*, Lanham, MD: Rowman & Littlefield, pp. 63–79), Figure 4.1.

¹²³ E.g. Knox, J. (2016). ‘Epistemic Mistrust: A Crucial Aspect of Mentalization in People with a History of Abuse?’, *British Journal of Psychotherapy*, 32(2): 226–236: ‘Epistemic trust is essentially a process of trust in the other person’s mind and therefore depends on mentalizing’ (p. 226).

trust—for instance, whether it can apply to the self, whether it can apply to feelings—this will hinder subsequent empirical work.

However, perhaps the greater urgency for theoretical specification lies in the fact that, in the past few years, Fonagy and colleagues have reoriented the goal of mentalization-based therapy. They have drawn on Bion's emphasis on curiosity and 'learning from experience' as essential for psychological health.¹²⁴ On this basis, mentalization is now considered only a means to an end: 'Mentalising in itself is only an intermediate step, not the ultimate therapeutic objective' because 'true and lasting improvement, we believe, rests on ... learning from experience beyond therapy'.¹²⁵ In general, 'mental disorder can be conceptualised as impairments in the capacity of the individual for social learning expressed in terms of epistemic trust'.¹²⁶ The question of the meaning and boundaries of epistemic trust and/or mistrust is therefore of the utmost importance. The way in which this question is answered will functionally delineate the goals of therapeutic work and the model of how to achieve them. It is also of great potential significance for conceptualizing the nature of mental illness, and how patients' symptoms should be grouped and assessed. A concern with epistemic trust cross-cuts existing diagnosis-focused approaches to mental health, suggesting potential revisions to this approach.

Diagnoses and the p-factor

To a large extent, the delivery, resourcing, and administration of clinical services for mental health—as well as the curriculum and structure of clinical training—have been oriented by discrete diagnoses. Diagnoses are functionally attractive for bureaucratic systems, because they permit judgements about prognosis and qualitative distinctions to be drawn about treatment need and delivery.¹²⁷ There have been long-standing critiques, however, of a diagnosis-centric approach to mental health. For instance, in his 1940 book, *Personality and Mental Illness*, John Bowlby described the assumption that mental illnesses 'were separate disease-entities each with its specific pathology' as wholly 'untenable'. He argued that more important than this or that symptom was 'the total mental state, a correct estimation of which will be the only reliable guide to prognosis'.¹²⁸ Following the publication of the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* by the American Psychiatric Association in 1980, Bowlby criticized the 'categorists' for their search 'for diagnostic criteria that distinguish the mental ill from the normal'. He urged collective critique of such a stance by 'those others who, like myself, believe continuity to be a more fruitful perspective'.¹²⁹ For instance, he argued that 'personality disorders' did not represent a discrete

¹²⁴ E.g. Bion, W. (1967). *Second Thoughts*, London: Karnac Books, p. 101: 'lack of progress in any direction must be attributed in part to the destruction of a capacity for curiosity and the consequent inability to learn'.

¹²⁵ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 33. See also Allison, E. and Fonagy, P. (2016). 'When is Truth Relevant?'. *Psychoanalytic Quarterly*, 85(2): 275–303: 'Mentalizing in this context is not an end in itself. Mentalizing is the catalyst that activates the effective ingredient of therapy: learning from experience' (p. 294).

¹²⁶ Fonagy, P. (2016). 'Epistemic trust and attachment: A fresh look at therapeutic processes in personality disorder'. AMBIT Conference, 2 July. Accessed at: <https://www.altrecht.nl/wp-content/uploads/2016/11/Ambit-Conference-Fonagy.pdf>.

¹²⁷ Jablensky, A. (2016). 'Psychiatric Classifications: Validity and Utility'. *World Psychiatry*, 15(1): 26–31.

¹²⁸ Bowlby, J. (1940). *Personality and Mental Illness*, London: Kegan Paul, p. 5.

¹²⁹ Bowlby, J. (1983). 'Darwin: Psychiatry and Developmental Psychology', Contribution to a Symposium on Darwin and Psychology held at the conference of the British Psychological Society, December. London: Wellcome Collection, PP/BOW/E.3/132. See Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press, Chapter 2.

kind of illness but an intensified and persistent form of ordinary difficulties humans may experience in recognizing, contributing to, and making use of the good that can come from social relationships. They ‘reflect an individual’s impaired ability to recognise suitable and willing figures and/or an impaired ability to collaborate in rewarding relationships with any such figure when found.’¹³⁰

Anna Freud was also a critic of diagnosis-focused mental health practice. In her essay, ‘The Symptomatology of Childhood’, she argued that ‘symptoms are no more than symbols, to be taken merely as indications that some mental turmoil is taking place’. Excessive emphasis should not be given to any individual symptom or set of symptoms, because ‘their place may be taken almost instantaneously by other pathological formations which, although overtly different, express the same latent content and may be no less aggravating for the individual’s life.’¹³¹ Part of Freud’s concern was that exactly the same set of symptoms, corresponding to a particular diagnostic profile, could reflect utterly different psychological processes, and vary in their clinical implications:

Symptoms may be no more than the child’s answer to some developmental stress and as such transitory, i.e., liable to pass away together with the maturational phase which has given rise to them. Or symptoms may represent a permanent countercathexis against some threatening drive derivative and as such be crippling to further development. Or symptoms, though pathological in origin, may nevertheless be ego-syntonic, and merged with the structure of the child’s personality to a degree which makes it difficult to distinguish between such manifestations as outward evidence of ongoing pathological involvement or as more or less normal, stable features of the individual’s character.¹³²

Freud thus advocated redirection of attention away from diagnostic categories and towards mental turmoil conceptualized as a latent factor. Within her clinic (latterly the Anna Freud Centre) she ‘developed a system of five diagnostic categories for classifying children’s psychopathology: category 1 (essentially normal), 2 (transient symptomatology or developmental strain), 3 (neurotic, with “permanent regressions, fixations and symptom formation”), 4 (atypical, distorted personality development), and 5 (destructive processes disrupting mental growth).’¹³³ This quasi-interval scale was among the outcome measures used in the Anna Freud Centre retrospective study (see Chapter 2). Another was the Hampstead Child Adaptation Measure developed by Target and Fonagy, in which ‘the general adjustment of a child’ was measured on a 100-point scale.¹³⁴

For their part, Fonagy and colleagues have acknowledged the pragmatic functions of diagnostic-based clinical practice and have argued against the abandonment of diagnoses. For instance, Fonagy was a co-author on a public letter advocating that diagnoses

¹³⁰ Bowlby, J. ([1970] 1979). ‘Self-Reliance and Some Conditions that Promote it’, in *The Making and Breaking of Affectional Bonds*, London: Routledge, pp. 124–149, p. 126.

¹³¹ Freud, A. (1970). ‘The Symptomatology of Childhood: A Preliminary Attempt at Classification’. *The Psychoanalytic Study of the Child*, 25(1):19–41, pp. 19–20.

¹³² Ibid. 21.

¹³³ Fonagy, P. and Target, M. (1994). ‘Who is Helped by Child Psychoanalysis? A Sample Study of Disruptive Children, from the Anna Freud Centre Retrospective Investigation’. *Bulletin of the Anna Freud Centre*, 17(4): 291–315, p. 296.

¹³⁴ Ibid. Recently, the AMBIT Adolescent Integrative Measure has been developed as an abbreviated version of the Hampstead Child Adaptation Measure. Fuggle, P., Bevington, D., Cracknell, L., Hanley, J., Hare, S., Lincoln, J., ... and Zlotowitz, S. (2015). ‘The Adolescent Mentalization-Based Integrative Treatment (AMBIT) Approach to Outcome Evaluation and Manualization: Adopting a Learning Organization Approach’. *Clinical Child Psychology and Psychiatry*, 20(3): 419–435.

can facilitate understanding the ‘interrelation of psychological processes in an individual patient’, by offering a picture of how symptoms may cohere.¹³⁵ Bateman and Fonagy have specifically defended the diagnosis of BPD, despite the fact that the category may be stigmatizing, and the fact that BPD is probably better conceptualized as a spectrum rather than a distinct clinical category.¹³⁶ Fonagy would prefer a label such as ‘persistent distress disorder’ rather than BPD.¹³⁷ But whatever the name he regards, on balance, some label as necessary to identify a group of patients with relatively distinct needs:

Mental health professionals have expressed a great deal of anxiety about giving a patient a diagnosis of personality disorder. Fears are rightly expressed about pejorative overtones, judgemental attitudes, blaming the patient, attacking the very ‘soul’ of the individual . . . Despite these potential drawbacks, we firmly believe that it is both necessary and constructive.¹³⁸

Discussing use of the BPD diagnosis for adolescents, Fonagy has argued that ‘it is time for clinicians to stop hanging on to outdated ideas about the alleged impossibility of diagnostics, saying that the diagnosis may be iatrogenic, that the disorder is “only a life phase” that “they will grow out of” or that there is insufficient data. These are all unsubstantiated concerns that do not work in the interest of young patients.’¹³⁹

To take another example: Fonagy has been willing to use the term ‘gender dysphoria’, because the diagnosis can help people questioning their gender identity to access specialist services. However, he has simultaneously expressed concern that ‘if we label gender dysphoria a mental disorder, and that disorder is stigmatizing, then that can be unhelpful and may actually reduce their willingness to explore the complexity of their feelings.’¹⁴⁰ Against those who characterize transgender as an identity necessitating swift movement to medical interventions, Fonagy has characterized anyone professing total certainty about gender identity as likely to be in a non-mentalizing state. Their statements about themselves should therefore be treated with a degree of doubt by clinicians. However, Fonagy equally described anyone who thinks that gender identity can be reduced to chromosomal differences to equally be in a non-mentalizing state (presumably teleological mode).¹⁴¹ There are people, both children

¹³⁵ Shedler, J., Beck, A. T., Fonagy, P., Gabbard, G. O., Kernberg, O., Michels, R., and Westen, D. (2011). ‘Response to Skodol Letter’. *American Journal of Psychiatry*, 168(1): 97–98, p. 97. See also Maldonado-Durán, M., Helmig, L., Moody, C., Fonagy, P., Fulz, J., Lartigue, T., . . . Glinka, J. (2003). ‘The Zero-to-Three Diagnostic Classification in an Infant Mental Health Clinic: Its Usefulness and Challenges’. *Infant Mental Health Journal*, 24(4): 378–397.

¹³⁶ See also Hutsebaut, J., Debbané, M., and Sharp, C. (2020). ‘Designing a range of mentalizing interventions for young people using a clinical staging approach to borderline pathology’. *Borderline Personality Disorder and Emotion Dysregulation*, 7(1): 1–10.

¹³⁷ Kirby, T. (2019). ‘Peter Fonagy—Battling the Enemy of Loneliness’. *The Lancet Psychiatry*, 6(12): 987.

¹³⁸ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press.

¹³⁹ Fonagy, P. (2015). ‘Peter Fonagy and the Undermining of Old Ideas on Personality Disorder’. Accessed at: <https://www.escap.eu/resources/resource-centre-disorders/undermining-of-old-ideas-on-personality-disorder>.

¹⁴⁰ Interview with Fonagy in Steidinger, S. (2018). ‘Trans-Actions: An Exploration of Gender Dysphoria’. Accessed at: <https://vimeo.com/285555219>: ‘Enhancing an individual’s capacity to represent their subjective experience in a more coherent and more fluid way would aid the resolution of gender identity issues in a child or young adult above other approaches that are more “radical”, shall we say. So I would strongly wish to support an initiative that aimed to enhance the capacity of an individual to conceptualize themselves and their subjective experience of their gender in a more complex way as a first stop in their journey to resolving gender dysphoria.’

¹⁴¹ See also Fonagy, P. and Allison, E. (2015). ‘A Scientific Theory of Homosexuality for Psychoanalysis’, in A. Lemma and P. E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, London: Routledge, pp. 125–137: ‘Drescher (2002) refers to this habit of thought as “binary thinking” . . . From a neuroscientific perspective, we might describe it as a result of failure of the mentalising capacity’ (p. 125).

and adults, who will need specialist services to respond to discrepancies between their gender identity and societal gender norms. In some cases, these discrepancies may stem from non-mentalizing, which is why a chance to talk to a psychological professional should be a first step. However, in other cases, discrepancies may remain for individuals who are mentalizing well, and then medical intervention may be helpful. While gender identity, like all self-representation, is always a confabulation for Fonagy, if a diagnosis like gender dysphoria can be made non-stigmatizing, then it will have value in allowing individuals to access specialist psychological and medical services.

In the case of BPD and gender dysphoria, Fonagy can be seen defending diagnostic categories. However, like Bowlby and Anna Freud, Fonagy and colleagues have more generally come down against diagnosis-focused mental health practice.¹⁴² In light of their account of non-mentalizing and epistemic mistrust as transdiagnostic risk factors, they have expressed concern at the limitations of a category-based system. They regard this system as neglecting five features of mental health that are of particular consequence for theoretical and clinical work, as well as health policy:

1. Continuities between ordinary difficulties and mental illness.
2. The interaction and patterns of co-occurrence between symptoms separated into discrete diagnoses.
3. The variability of potential symptoms within many diagnoses which may mask different aetiologies and trajectories.
4. The variability in day-to-day functioning for individuals with the same diagnosis—or with no diagnosis.
5. Factors that specifically contribute to thriving and resistance to mental illness, as distinct from vulnerability or risk factors.¹⁴³

¹⁴² Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press. 'Even if we do not believe in categorical diagnostic systems, someone else in the mental health service is likely to have given the patient a diagnosis' (p. 149).

¹⁴³ Chiesa, M., Cirasola, A., Williams, R., Nassisi, V., and Fonagy, P. (2017). 'Categorical and Dimensional Approaches in the Evaluation of the Relationship between Attachment and Personality Disorders: An Empirical Study', *Attachment & Human Development*, 19(2): 151–169. See also St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). 'Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults', *PLoS One*, 12(4): e0175381; Bergen, L., and Grimes, T. (1999). 'The Reification of Normalcy', *Journal of Health Communication*, 4(3): 211–226: 'Because normalcy, when it is clinically considered, actually implies a wide variation in mental states (DSM-IV American Psychiatric Association, 1994), we believe that one must assume that "normal" children can vary greatly in their reaction and susceptibility to violent television programming. The reason many ill or prodromal children (that is, children who do not manifest DSM-IV classifiable behaviors but who show early symptoms of potentially classifiable mental disorders) may be viewed as normal is because most lay people, perhaps communication researchers among them, consider normalcy to be characterized by the ability to operate within the world satisfactorily enough to "get along," to interact passably or better with one's friends, family, and colleagues (P. Fonagy, personal communication, January 8, 1998). This definition of normalcy does not preclude the presence of psychopathology, in either a clinical or prodromal stage (P. Fonagy, personal communication, January 8, 1998). Indeed, most of the people in the population who are ill are undiagnosed and are generally treated as being normal' (p. 216). Most recently, see Wendt, L. P., Wright, A. G., Pilkonis, P. A., Nolte, T., Fonagy, P., Montague, P. R., ... and Zimmermann, J. (2019). 'The latent structure of interpersonal problems: Validity of dimensional, categorical and hybrid models', *Journal of Abnormal Psychology*, 128(8): 823–839. Fonagy has also repeatedly drawn attention to lifetime prevalence data for mental illness as suggesting deficits in the assumption of health in the absence of a diagnosis: Schaefer, J. D., Caspi, A., Belsky, D. W., Harrington, H., Houts, R., Horwood, L. J., ... and Moffitt, T. E. (2017). 'Enduring Mental Health: Prevalence and Prediction', *Journal of Abnormal Psychology*, 126(2): 212–224: 'Only 17% of repeatedly assessed Study members managed to reach midlife (age 38) without experiencing the psychiatric symptoms and resulting functional impairment necessary to meet criteria for the diagnosis of a mental disorder' (p. 220).

Fonagy and colleagues have also highlighted findings by Goldberg and colleagues, which show that, while there is much individual variation, on average there is a small decline in clinical effectiveness for every year of experience a therapist accrues.¹⁴⁴ Fonagy and colleagues conclude that these findings ‘might arise from the fact that less experienced therapists are more able to see the individual patient in all his or her subjective complexity rather than as a “walking diagnostic prototype”’.¹⁴⁵ Such prototyping undermines the clinician’s recognition of the specificity of their patient and their mental states, which would form the basis for prompting epistemic trust by the patient. Fonagy and colleagues have acknowledged that the spread of manualized therapy, a development they otherwise regard as positive and have helped support, may nonetheless have facilitated this prototyping process.¹⁴⁶ This has prompted the use of web-based user-adaptable treatment manuals at the Anna Freud Centre for developing mentalizing teams within organizations,¹⁴⁷ though conventional manuals currently remain used in most MBT.

Fonagy and colleagues have seen support for their criticisms of diagnosis-focused mental health practice in work by Caspi and colleagues on the ‘p-factor’, a transdiagnostic latent construct in the formation of mental health symptoms.¹⁴⁸ As we have seen, precursors of the idea of transdiagnostic latent construct can be found already in the work of Bowlby and Anna Freud. In fact, already in 1946, the outgoing President of the British Psychoanalytic Society, Ernest Jones, had theorized in his valedictory address that later researchers would discover a general psychopathology factor equivalent to the general intelligence factor assessed as IQ.¹⁴⁹ Yet, until recently, this hypothesis has lacked empirical support. In 2014, Caspi and collaborators reported from a factor analysis of mental health symptoms shown by participants in the Dunedin Multidisciplinary Health and Development Study. The diverse symptoms initially clustered into three higher order factors: internalizing symptoms, externalizing symptoms, and disorders of thought. However, more variance was accounted for by a model assuming one general psychopathology dimension. Caspi and colleagues termed this the ‘p-factor’. They found that the p-factor was associated with three character

¹⁴⁴ Goldberg, S., Rousmaniere, T., Miller, S., Whipple, J., Nielsen, S., Hoyt, W., and Wampold, B. (2016). ‘Do Psychotherapists Improve With Time and Experience? A Longitudinal Analysis of Outcomes in a Clinical Setting.’ *Journal of Counseling Psychology*, 63(1): 1–11.

¹⁴⁵ Fonagy, P., Campbell, C., and Allison, E. (2019). ‘Therapeutic Models’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 171. Fonagy has drawn comparison with ideas from Zen Buddhism here, citing, ‘In the beginner’s mind there are many possibilities, in the expert’s mind there are few.’ Suzuki, S. (1970). *Zen Mind, Beginner’s Mind*, New York: Weatherall, p. 21. Fonagy, P. (2019). ‘Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience.’ GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast/>.

¹⁴⁶ Ibid.

¹⁴⁷ Bevington, D. (2020). ‘Towards a Learning Stance In Teams: Developing a Community of Practice to Capture and Disseminate what Works for Whom’, in Ilana Crome and Richard Williams (eds), *Substance Misuse and Young People: Critical Issues*, London: Routledge, pp. 450–466: ‘each team inherits all the generic AMBIT content as the starting point for its local manual. Teams are trained to engage in regular brief discussions about elements of practice, described in more detail below, that end in their producing key bullet points that can be owned by the team (manualising is thus a collective activity, never a single worker’s ideas). Teams either overwrite their own local versions of pages inherited from AMBIT or can add new material alongside this inherited content. Thus, local manuals are co-constructions that draw together centrally curated evidence-based practice and locally generated practice-based evidence ... The process of a team manualising its local practice is seen as analogous to the group mentalising itself because it inquires and reflects on why and how we do things in this way’ (p. 459).

¹⁴⁸ Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... and Moffitt, T. E. (2014). ‘The P Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders?’. *Clinical Psychological Science*, 2(2): 119–137. See also Laceulle, O. M., Vollebergh, W. A., and Ormel, J. (2015). ‘The Structure of Psychopathology in Adolescence: Replication of a General Psychopathology Factor in the TRAILS Study’. *Clinical Psychological Science*, 3(6): 850–860.

¹⁴⁹ Jones, E. (1946). ‘A Valedictory Address’. *The International Journal of Psychoanalysis*, 27: 7–12.

traits: low Agreeableness, low Conscientiousness, and high Neuroticism. The researchers found that child maltreatment was associated with the p-factor, but not with the lower-level factors once the p-factor was included in the model.

In 2015, Fonagy and Campbell acknowledged that there may be diverse forms of mental illness with particular properties, and even predictable profiles of symptoms that give the impression of relative distinct disorders. Nonetheless, Fonagy and Campbell argued, ambitiously, that ‘the p factor is a measurement of epistemic trust’.¹⁵⁰ In a later statement, Fonagy has been more cautious: ‘it might be helpful to consider the p factor as a proxy for impairments in epistemic trust’.¹⁵¹ Whether strongly or more tentatively presented, in the perspective offered by Fonagy and colleagues, the initiation and maintenance of symptoms can be accounted for in terms of epistemic mistrust and its relationship with non-mentalizing modes; and their prevention and reduction can be accounted for in terms of epistemic trust and its relationship with mentalization. Fonagy and colleagues argued that the character traits of low Agreeableness, low Conscientiousness, and high Neuroticism identified by Caspi and collaborators should be regarded as reflecting ‘emotion dysregulation, impulsivity and social dysfunction’, the key symptoms of BPD.¹⁵² While these features may be misrecognized as aspects of a person’s personality, Fonagy and colleagues proposed that all three should be regarded as stemming from dysfunctions in epistemic trust. For instance, epistemic mistrust leads to non-mentalizing modes, which in turn prompt impulsive action because the potential for reconsideration is taken offline.¹⁵³

Collaborators of Fonagy have pursued this perspective through empirical studies. Sharp and colleagues conducted a factor analysis of the symptoms of 966 inpatients in a paper published in 2015. On the one hand, there were some distinguishing features for specific forms of personality disorder. Disregard for safety, aggression, and lack of remorse seemed to form a profile aligned with the antisocial personality disorder (ASPD) diagnosis. Odd behaviour and beliefs aligned with the schizotypal personality disorder diagnosis. Grandiosity and need for admiration aligned with the diagnosis of narcissistic personality disorder. Moral inflexibility aligned with the obsessive–compulsive personality disorder. And preoccupation with rejection aligned with avoidant personality disorder. On the other hand, most of the variance in symptoms could be accounted for on the assumption of a general psychopathology

¹⁵⁰ Fonagy, P. and Campbell, C. (2015). ‘Bad Blood Revisited: Attachment and Psychoanalysis, 2015’. *British Journal of Psychotherapy*, 31(2): 229–250, p. 243

¹⁵¹ Fonagy, P. (2016). ‘The Role of Attachment, Epistemic Trust and Resilience in Personality Disorder: A Trans-Theoretical Reformulation’. DMM News, 26 September. Accessed at: <http://www.iasa-dmm.org/images/uploads/DMM%20%2322%20Sept%2016%20English.pdf>. The most recent formulation is also more cautious: ‘many manifestations of mental disorder may be underpinned by an inability to benefit from social communication due to epistemic mistrust’. Fonagy, P., Allison, E., and Campbell, C. (2019). ‘Mentalising, Resilience and Epistemic Trust’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 71.

¹⁵² Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). ‘What We have Changed our Minds About: Part 1. Borderline Personality Disorder as a Limitation of Resilience’. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 11.

¹⁵³ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: ‘We understand impulsivity in terms of imbalance among the poles of mentalising: it involves a heavy emphasis on the automatic pole of the automatic-controlled dimension ... What reflection there is will most likely be disconnected from reality [pretend mode] ... I will entangle myself and my clinician in extended analyses, but offer little compelling evidence for any of my assertions. I urgently seek validation for my view, but even when it is forthcoming, it is meaningless because I am simultaneously aware that I made up my explanation; thus confirming or elaborating it only increases my sense of emptiness and meaninglessness’ (p. 44); ‘Impulsivity can also result from the teleological stance ... for such a patient, what does seem to help are physical actions that make him/her feel real’ (p. 45).

factor.¹⁵⁴ Sharp and colleagues were interested that no specific profile emerged for BPD. Instead, the items associated with BPD fell almost exclusively with the general factor.¹⁵⁵ This included identity disturbance, feelings of emptiness, self-harm, instability of social relationships, and social actions motivated by a fear of abandonment. From this, Sharp and colleagues concluded that the criteria for BPD do not reflect a 'type' of personality disorder, but 'best reflect general impairments'.¹⁵⁶ This appears to offer empirical support for a position first put forward by Fonagy's friend Otto Kernberg in the 1980s, that BPD is not a discrete form of mental illness but an indication of general psychosocial difficulties.¹⁵⁷ On this basis of both theory and empirical findings, Bateman and Fonagy offered the clinical observation that, when symptoms of various forms of personality disorder are present, 'treatment of BPD is required if comorbid disorder is to be treated', because it is the general difficulties characteristic of BPD that are the more basic, and fundamentally causal.¹⁵⁸

In another study, also published in 2015, Patalay, Fonagy, and colleagues conducted a factor analysis of self-reported symptoms in a community-based sample of 23,477 early adolescents, with follow-up data after one year for 10,270 participants.¹⁵⁹ They found that most of the variance in symptoms in their sample could be accounted for in terms of a general p-factor, with lower-level dimensions of internalizing and externalizing symptoms making a small additional contribution (see Figure 7.4). Without the general p-factor taken into account, internalizing and externalizing symptoms were clearly correlated ($r = .45$). However, after removing variance associated with the p-factor, internalizing and externalizing symptoms were negatively correlated. The associations between female gender and internalizing symptoms, and between male gender and externalizing symptoms, also increased. This suggests that the p-factor may have been masking the extent of the gender specificity of these symptoms in previous epidemiological research on mental health.

In terms of prediction, the general p-factor had a strong association with mental health symptoms a year on, with the internalizing and externalizing dimensions comparatively less effective as predictors. Patalay, Fonagy, and colleagues claimed that diagnosis-focused mental health practice focuses attention on specific pathways to specific problems, ignoring the general factors that provide the motive force behind mental illness. In particular, the researchers proposed that diagnosis-focused mental health practice fails to consider the way

¹⁵⁴ An exception was symptoms of narcissistic personality disorder, which turned out to be quite distinct: Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., and Clark, L. A. (2015). 'The Structure of Personality Pathology: Both General ('g') and Specific ('s') Factors?'. *Journal of Abnormal Psychology*, 124(2): 387–398: 'Narcissistic PD criteria's average loading on the general factor was rather weak ($r = .31$). What implications does this have for how we conceptualize, for example, narcissistic criteria and traits?' (p. 396). However, the distinct qualities of symptoms of narcissistic personality disorder have not been found to replicate: Wright, A. G., Hopwood, C., Skodol, A., and Morey, L. (2016). 'Longitudinal Validation of General and Specific Structural Features of Personality Pathology'. *Journal of Abnormal Psychology*, 125(8): 1120–1134.

¹⁵⁵ These findings were subsequently replicated in Ibid. Here the relationship between BPD and the p-factor was even stronger than in the study by Sharp and colleagues.

¹⁵⁶ Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., and Clark, L. A. (2015). 'The Structure of Personality Pathology: Both General ('g') and Specific ('s') Factors?'. *Journal of Abnormal Psychology*, 124(2): 387–398, p. 394.

¹⁵⁷ Kernberg, O. F. (1984). *Severe Personality Disorders: Psychotherapeutic Strategies*. New Haven, CT: Yale University Press.

¹⁵⁸ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'treatment of BPD is required if comorbid disorder is to be treated; however, treating the comorbid disorder does not improve BPD' (p. 82). See also Fonagy, P., Luyten, P., and Bateman, A. (2017). 'Treating Borderline Personality Disorder with Psychotherapy: Where Do We Go From Here?'. *JAMA Psychiatry*, 74(4): 316–317.

¹⁵⁹ Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., and Wolpert, M. (2015). 'A General Psychopathology Factor in Early Adolescence'. *The British Journal of Psychiatry*, 207(1): 15–22.

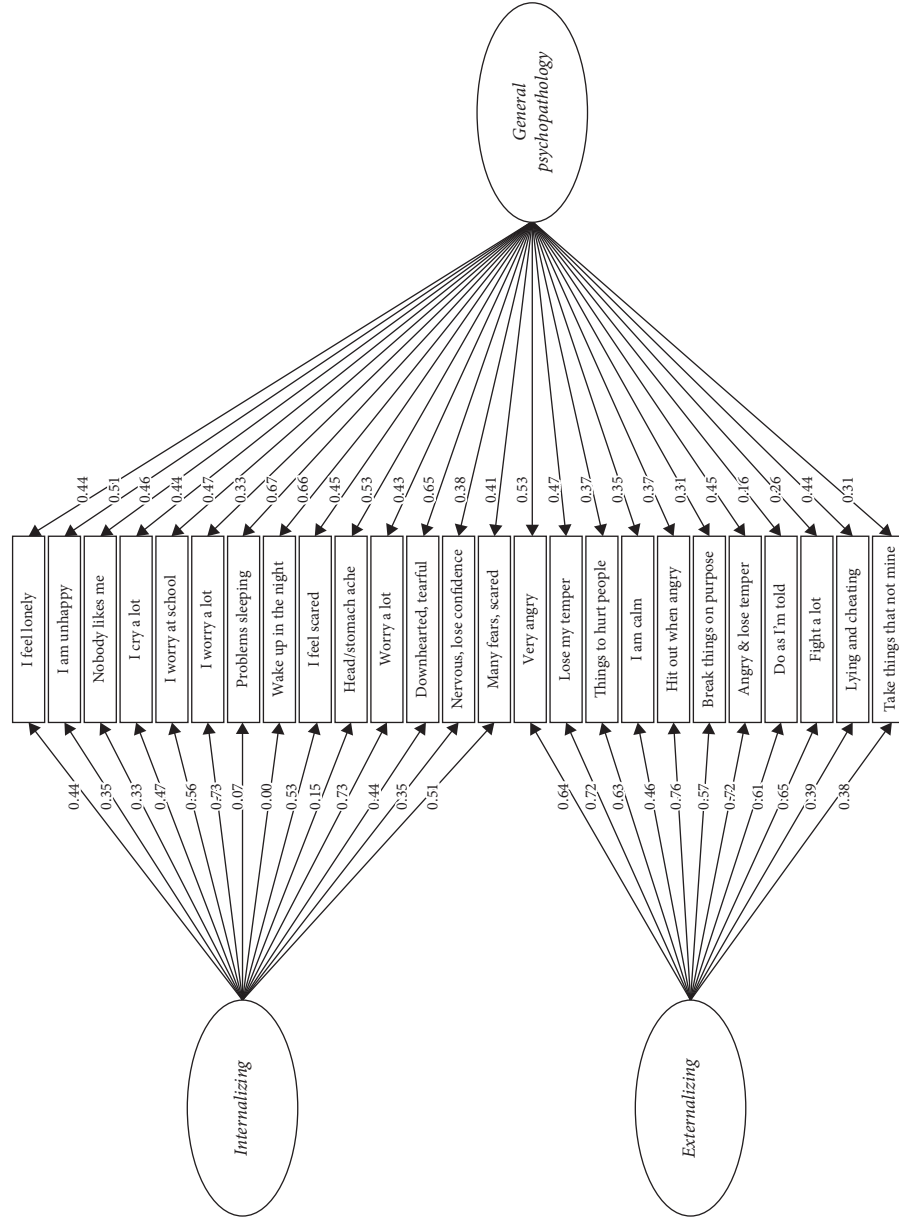


Figure 7.4 Bi-factor model with the item loadings onto the internalizing and externalizing dimensions and the general psychopathology bi-factor.

Source: Reproduced from Praveetha Patalay et al., 'A General Psychopathology Factor in Early Adolescence', *British Journal of Psychiatry*, 207(1): 15–22, Figure 1, DOI: <https://doi.org/10.1192/bjp.bp.114.149591> Copyright © Royal College of Psychiatrists, 2015.

that mental illnesses of various kinds may be prompted and maintained by the breakdown of processes that usually offer protection, above all epistemic trust and, stemming from it, the capacity to make use of social relationships for responding effectively to adversities. The prediction from the general p-factor more than the internalizing and externalizing dimensions in the study is a finding with multiple ramifications. For instance, it implies that trials of psychological interventions may be radically underestimating effectiveness, if diagnoses are used as outcome measures rather than a general p-factor.

A later study by St Clair, Neufeld, Jones, Fonagy, and colleagues reported from a study of the self-reported symptoms of 2,257 volunteer participants aged between 14 and 25.¹⁶⁰ Again, a general latent factor accounted for much of the variance. There were also five lower-level factors: self-confidence, antisocial behaviour, worry and fear, aberrant thinking (including obsessional/compulsive and psychotic-like experience items), and mood. One important finding was that states of anxiety and depression,¹⁶¹ as well as items related to somatic functioning (e.g. 'moved and walked more slowly than usual', 'didn't sleep as well as usual'), loaded exclusively on the general factor and not on any of the lower-level factors. This aligns with claims made by Luyten, Fonagy, and colleagues that somatic disorders should be regarded as expectable features of the continuum of mental distress, rather than a discrete diagnostic category.¹⁶²

A second finding of particular theoretical importance from the St Clair study was that agreement or disagreement with self-report items about negative aspects of mental health only offered good discrimination at the high end of the general and specific factors. By contrast, agreement or disagreement with statements characterizing positive aspects of mental health offered discrimination at both ends of the spectrum. St Clair and colleagues argued that this underlines the necessity of concern with positive factors that protect against mental illness in understanding why people suffer more or less mental illness. This finding likewise aligns well with mentalization and epistemic trust situated by Fonagy and colleagues as transdiagnostic sources of resilience, which would otherwise be missed through a focus only on diagnostic pathology. However, it also suggests that risk factors may not simply be the inverse of protective factors: each may have distinct correlates. As we saw earlier, epistemic mistrust is not simply the opposite of epistemic trust, and they play rather different roles in Fonagy and colleagues' developmental model. It may likewise be wondered whether mentalization is simply the opposite of non-mentalizing, or whether mentalization and non-mentalizing both respectively have distinct qualities and correlates that theory and measurement should seek to capture.¹⁶³

¹⁶⁰ St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). 'Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults'. *PLoS One*, 12(4): e0175381.

¹⁶¹ St Clair and colleagues identified a lower-level factor of long-term, stable negative mood, which seemed to have different properties from more state-like components of depression, the latter loading exclusively on the general p-factor. St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). 'Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults'. *PLoS One*, 12(4): e0175381.

¹⁶² Luyten, P., Van Assche, L., Kadriu, F., Krans, J., Claes, L., and Fonagy, P. (2017). 'Other Disorders often Associated with Psychological Trauma', in C. Dalenbergh, S. Gold, and J. Cook (eds), *APA Handbook of Trauma Psychology*. Volume 1: Foundations in Knowledge, Washington, DC: American Psychological Association, pp. 243–280, p. 245.

¹⁶³ The potential for mentalization to contribute to broaden-and-build cycles, but not the absence of non-mentalizing, has been discussed recently by Bo, S., Bateman, A., and Kongerslev, M. T. (2019). 'Mentalization-Based Group Therapy for Adolescents with Avoidant Personality Disorder: Adaptations and Findings from a Practice-Based Pilot Evaluation'. *Journal of Infant, Child, and Adolescent Psychotherapy*, 18(3): 249–262.

A further notable finding from the study was that while psychotic experiences loaded in part with the aberrant thinking lower-level factor, they were also strongly associated with the general psychopathology factor. St Clair and colleagues interpreted the findings as showing that psychotic symptoms should also be regarded as expectable features of the continuum of the p-factor, perhaps characterizing especially its most comorbid and/or its most severe end, rather than a discrete diagnostic category.¹⁶⁴ Some psychotic symptoms, as much as other mental health symptoms within the overall p-factor, may be sequelae of vulnerabilities such as early trauma.¹⁶⁵ Developing ideas initially formulated by Fonagy in the treatment of Mr S (see Chapter 1), Debbané, Brent, Fonagy, and colleagues have proposed that, like other mental health difficulties, vulnerabilities to psychotic experiences will be exacerbated by emotion dysregulation and biases in interpersonal attribution characteristic of the p-factor. These will then remain uncorrected—and treatment will prove less effective—to the extent that modes of non-mentalization are used intensively, and epistemic vigilance locked in place.¹⁶⁶ Teleological mode can hold individuals to a single way of interpreting others' behaviour. Psychic equivalence can make passing thoughts or feelings seem characteristic of all reality. Pretend mode can sustain states of dissociation or hallucination, with hypermentalization contributing to improbable over-interpretation of others' behaviour and motives. Externalization of the alien self can make an individual's own negative feelings appear to be coming from outside. Fonagy and colleagues have regarded symptoms of psychosis either as a facet of the overall p-factor, or else as a subfactor on a level with internalizing and externalizing. In either case, they have conceptualized mentalization as protective against psychosis, and non-mentalization as predisposing and potentially exacerbating psychosis. However, there remain significant ambiguities regarding this claim. Not least, the precise basis of the conceptual distinction between pretend mode and psychosis remains under-elaborated: is pretend mode an independent contributor to (some forms of) psychosis, a component of (some forms of) psychosis, solely a reinforcer of psychotic symptoms, among the consequence of (some forms of) psychosis, or all four?

A final intriguing finding from the St Clair study was that higher scores on the p-factor, but not the specific factors, were associated generally with attending treatment for mental ill health. The exception here was that both the p-factor and the mood lower-level factor predicted attending treatment for depression. However, participants who *only* reported high symptoms of the p-factor and not the symptoms of any of the lower-level factors were no more likely to be receiving mental health services than individuals with few or no symptoms. There are various interpretations that could be offered of this finding. One would be that there is significant unmet need among the population of individuals with only symptoms of the p-factor. St Clair and colleagues offered a very different interpretation, proposing

¹⁶⁴ See also Stochl, J., Khandaker, G. M., Lewis, G., Perez, J., Goodyer, I. M., Zammit, S., ... and Jones, P. B. (2015). 'Mood, anxiety and psychotic phenomena measure a common psychopathological factor'. *Psychological Medicine*, 45(7): 1483–1493.

¹⁶⁵ See also van Nierop, M., Viechtbauer, W., Gunther, N., Van Zelst, C., De Graaf, R., Ten Have, M., ... and Outcome of Psychosis (GROUP) investigators (2015). 'Childhood Trauma is Associated with a Specific Admixture of Affective, Anxiety, and Psychosis Symptoms Cutting across Traditional Diagnostic Boundaries'. *Psychological Medicine*, 45(6): 1277–1288.

¹⁶⁶ Debbané, M., Salaminios, G., Luyten, P., Badoud, D., Armando, M., Solida Tozzi, A., Fonagy, P. and Brent, B. K. (2016). 'Attachment, Neurobiology, and Mentalizing along the Psychosis Continuum'. *Frontiers in Human Neuroscience*, 10: 406; Debbané, M. and Toffel, E. (2019). 'Mentalizing Through the Early Stages of the Psychosis Continuum', in João G. Pereira, Jorge de Almeida Gonçalves, and Valeria Bizzari (eds), *The Neurobiology-Psychotherapy-Pharmacology Intervention Triangle*, Wilmington, DE: Vernon Press, pp. 141–160; Debbané, M. and Bateman, A. (2019). 'Psychosis', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 417–429.

that these individuals have less need of clinical services. They therefore questioned the widespread use of measures of anxiety and depression in screening in primary and secondary care settings, because their interpretation was that patients with both specific and general mental health symptoms were the ones with the most need. Evidence in favour of this interpretation is that elevations of the p-factor alone, in the absence of any of the lower-level factors, was also not associated with an increase in use of alcohol, cannabis, illegal drugs, or non-suicidal self-injury. By contrast, the likelihood of report of these behaviours was elevated for participants scoring highly on both the p-factor and the anti-social lower-level factor, when compared with just those participants scoring highly on the anti-social factor.¹⁶⁷

In a further study on the same sample, Polek and colleagues found that the p-factor contained high loadings on various sets of items.¹⁶⁸ One set were directly suggestive of epistemic mistrust. This included endorsement of the item 'I feel I have to be on my guard even with friends', as well as low agreeableness and high antagonism. A second set indicated a perception of the views and norms of others as lacking relevance to the self. This included items such as 'I do not care if I get into trouble' and 'I do not care about doing things well'. Another set were items suggestive of weak integration of self-representation, of emotion, and of social relationships, such as affirmative answers to 'I am an odd, unusual person' and 'Do your moods change unpredictably?'. This was reflected in strong loading on self-report of feelings of social isolation, such as affirmative responses to the item: 'Do you feel that you cannot get close to people?' A further set were items from a measure of impulsivity, specifically those related to difficulties managing attention such as 'I concentrate easily' (reversed) and 'I don't pay attention', as well as reporting acting without thinking about the consequences.¹⁶⁹ Such findings are well aligned with the account by Fonagy and colleagues of close relations between epistemic trust, difficulties in social relationships, difficulties in the modulation of anger and sadness, instability and incoherence in self-representations, and disruption of attentional processes.

Recently, Gibbon, Nolte, and Fonagy examined the factor structure of self-report data from a large study. There were 338 participants from two sources: either patients with BPD, recruited from clinical services in London; or patients with ASPD recruited from probation services. Active psychosis was an exclusion criteria.¹⁷⁰ The study also included 167

¹⁶⁷ On suicidality as reflecting the p-factor, see Prabhu, G., Dolan, R. J., Bullmore, E. T., Fonagy, P., Stochl, J., Jones, P. B., and NSPN Consortium. (2020). 'How do the Prevalence and Relative Risk of Non-Suicidal Self-Injury and Suicidal Thoughts Vary across the Population Distribution of Common Mental Distress (the P-Factor)? Observational Analyses Replicated in Two Independent UK Cohorts of Young People'. *BMJ Open*, 10: e032494. Fonagy and Bateman have contrasted BPD and ASPD on several grounds. These contrasts include that 'people with ASPD are more likely to demonstrate over-control of their emotional states within well-structured, schematic attachment relationships, rather than the under-control in chaotic attachment relationships that are more commonly found in people with BPD. Their dyscontrol may be limited to volatile anger: Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 377–378.

¹⁶⁸ Polek, E., Jones, P. B., Fearon, P., Brodbeck, J., Moutoussis, M., Dolan, R., ... and NSPN Consortium. (2018). 'Personality Dimensions Emerging during Adolescence and Young Adulthood are Underpinned by a Single Latent Trait Indexing Impairment in Social Functioning'. *BMC Psychiatry*, 18(1): 23, online additional file, Table 1.

¹⁶⁹ Other strongly loading items were: 'Your emotions are shallow and fake'; 'You use or con other people to get what you want'; 'I seem very cold and uncaring to others'; 'People sometimes find me aloof and distant'; 'Some people find me a bit vague and elusive during a conversation'; 'I feel very uncomfortable in social situations involving unfamiliar people'.

¹⁷⁰ Gibbon, L., Nolte, T. and Fonagy, P. (2017). 'Modelling Axis I and Personality Disorder Symptomatology and its Associations with Childhood Trauma and Mentalising', in L. Gibbon, *An Interpretive Thematic Analysis of the P-Factor Literature and an Empirical Investigation of the Relationship between the P-Factor and Childhood Trauma and Reflective Function Childhood Trauma and Reflective Function*. Unpublished DClinPsy thesis, London: University College London. Accessed at: http://discovery.ucl.ac.uk/1574529/1/Thesis_final_volume1_Gibbon.pdf.

community volunteers to serve as controls. This study was the first to consider symptoms of personality disorder within the same analysis as other symptoms of mental ill health.¹⁷¹ The analysis found that the best fit for the data was a general p-factor plus four lower-level dimensions: internalizing, antisocial/hostility, thought disorder, and borderline. The symptoms that loaded most clearly on the general p-factor were interpersonal sensitivity, depression, anxiety, and psychoticism. Phobias, obsessive-compulsivity, post-traumatic stress, paranoid thoughts, and high scores on the 'Difficulties in Emotion Regulation Scale' were also, to an extent, characteristic of the general p-factor. This was much in line with the findings of Caspi and colleagues. The internalizing dimension was notable for having strong associations with the other lower-level dimensions. Symptoms of post-traumatic stress fell strongly with the internalizing dimension. In fact, there were significant indications that the internalizing dimension might be better modelled as part of the p-factor. By contrast, the antisocial/hostility lower-level dimension was notable for having few associations with the other lower-level dimensions, and items loaded only weakly on to the p-factor.

Gibbon, Nolte, and Fonagy found that self-report of childhood maltreatment was strongly associated with all four lower-level factors ($r_s = .53-.59$). But these associations were very substantially attenuated (to $r_s = .20-.28$) when the p-factor was included ($r = .55$), indicating that it was the associations with the p-factor that were driving the links to the lower-level factors. Self-report retrospective measures of childhood should not, however, be taken to be equivalent to the findings from prospective studies of maltreatment: a recent meta-analysis has shown that 52% of individuals with prospective observations of childhood maltreatment do not report it retrospectively, and 56% of individuals retrospectively reporting childhood maltreatment were not identified as maltreated in prospective studies.¹⁷² And McCrory and other colleagues at the Anna Freud Centre have specifically criticized the measure of childhood maltreatment used in the Gibbon study as overinclusive—for instance, identifying being called 'lazy' or 'stupid' by people in one's family—including siblings—as child maltreatment.¹⁷³

Gibbon and colleagues reported findings for reflective function, measured using the RFQ. For the certainty sub-scale (RFQc), associations with the four lower-level dimensions were moderate to strong ($r_s = .47-.55$). However, these associations mostly disappeared entirely once the p-factor was included ($r = .50$), with the exception of the borderline lower-level dimension which retained a moderate correlation ($r = .34$). For the uncertainty sub-scale (RFQu), associations with the four lower-level dimensions were remarkably strong ($r_s = .62-.70$). These associations were very substantially attenuated after inclusion of the p-factor.¹⁷⁴ The RFQu had an association with the p-factor of $r = .62$, which is an unusually large effect size for a psychological study, to the point that it could suggest overlap between the constructs.¹⁷⁵ Nonetheless, it should be pointed out that again the association between RFQu

¹⁷¹ A previous study did include other relevant symptoms, but not symptoms of BPD specifically, in the analysis: Hoertel, N., Franco, S., Wall, M. M., Oquendo, M. A., Kerridge, B. T., Limosin, F., and Blanco, C. (2015). 'Mental Disorders and Risk of Suicide Attempt: A National Prospective Study'. *Molecular Psychiatry*, 20(6): 718.

¹⁷² Baldwin, J. R., Reuben, A., Newbury, J. B., Danese, A. (2019). 'Agreement between Prospective and Retrospective Measures of Childhood Maltreatment: A Systematic Review and Meta-Analysis'. *JAMA Psychiatry*, 76(6): 584-593.

¹⁷³ McCrory, E. J., Gerin, M. I., and Viding, E. (2017). 'Annual Research Review: Childhood Maltreatment, Latent Vulnerability and the Shift to Preventative Psychiatry—The Contribution Of Functional Brain Imaging'. *Journal of Child Psychology and Psychiatry*, 58(4): 338-357.

¹⁷⁴ However, the internalizing and thought disorder dimensions remained barely significant correlates of RFQu even after taking into account the p-factor ($r_s = .20$ and $.17$ respectively).

¹⁷⁵ Funder, D. C. and Ozer, D. J. (2019). 'Evaluating Effect Size in Psychological Research: Sense and Nonsense'. *Advances in Methods and Practices in Psychological Science*, 2(2): 156-168.

and the borderline dimension was not entirely attenuated when the p-factor was included, but remained moderate ($r = .37$), and the antisocial/hostile dimension also retained a meaningful association ($r = .28$). The strongest association between the RFQu and individual symptoms was with items from the 'Difficulties in Emotion Regulation Scale' ($r = .69$) and emotional instability ($r = .61$). The strongest association between the RFQc and individual symptoms was with items from the 'Difficulties in Emotion Regulation Scale' ($r = .56$), items understood to signal problems with identity¹⁷⁶ ($r = .51$) and phobia ($r = .47$).

Overall, these findings offer clear support for the prediction by Fonagy and colleagues that the p-factor would be associated with non-mentalizing, though they align with other findings in recent years that generally suggest that RFQu has better associations with expectable correlates than RFQc. For instance, no association has been found for the RFQc with reported interpersonal difficulties or with the 'Difficulties in Emotion Regulation Scale', or with observer-reported measures of mentalizing such as the Parent Development Interview.¹⁷⁷ The findings also suggest potential qualifications to theory. Antisocial/hostile behaviour was only weakly associated with the p-factor, and retained a significant association with RFQu even after the p-factor was included in the model. Yet more strongly, excessive certainty and uncertainty regarding mental states was associated with borderline symptoms even when the p-factor was taken into account. This qualifies the implication that the high end of the p-factor = borderline = low reflective function. There appear to be aspects of BPD—and conduct problems to an extent—predicted by non-mentalizing that are distinct from the p-factor. However, these findings should be treated with caution because they are likely to have been affected by the specific recruitment of patients with BPD and ASPD, which may have bolstered the apparent independence of these lower-level factors. It might be, for instance, that the thought disorder lower-level factor would have retained its association with the RFQ if the sample had purposefully included a group of patients with psychotic and other anomalous experiences, rather than excluding participants with active psychosis. Or that the internalizing lower-level factor might have had more independence from the p-factor if the sample had included a group of patients with dysthymia.¹⁷⁸

¹⁷⁶ As measured with the Personality Assessment Inventory—Borderline Features. The 'identity problems' subscale comprises the following questions: 'My attitude about myself changes a lot'; 'Sometimes I feel terribly empty inside'; 'I worry a lot about other people leaving me'; 'I often wonder what I should do with my life'; 'I can't handle separation from those close to me very well'; and 'I don't get bored very easily' (reversed). Criticism of the psychometric qualities of the RFQc scale has been offered by Spitzer and colleagues. For instance they observe that 'Although RFQ_C is supposed to represent certainty about mental states, all items are geared towards a state of uncertainty with respect to their semantic content (e.g., item 1: "I find the thoughts of others confusing", or item 3: "When I get angry, I say things without really knowing why I am saying them") and are, ultimately, reversely scored. Thus, the certainty scale is based entirely on the denial of uncertainty'. Müller, S., Wendt, L. P., Spitzer, C., Masuhr, O., Back, S. N., and Zimmermann, J. (2020). A critical evaluation of the reflective functioning questionnaire. <https://psyarxiv.com/5rhme/>. Despite such problems with the scale, for recent evidence that parental RFQc at the start of family MBT can serve as a predictor of treatment outcome, see Jewell, T., Herle, M., Serpell, L., Eivors, A., Simic, M., Fonagy, P., and Eisler, I. (2020). 'Attachment and mentalization as predictors of outcome in family therapy for adolescent anorexia nervosa'. *PsyArXiv* pre-print, <https://psyarxiv.com/fgytkl/>

¹⁷⁷ Euler, S., Nolte, T., Constantinou, M., Griem, J., Montague, P. R., Fonagy, P., and Personality and Mood Disorders Research Network (2019). 'Interpersonal Problems in Borderline Personality Disorder: Associations with Mentalizing, Emotion Regulation, and Impulsiveness'. *Journal of Personality Disorders*, Early View; Handeland, T. B., Kristiansen, V. R., Lau, B., Håkansson, U., and Øie, M. G. (2019). 'High Degree of Uncertain Reflective Functioning in Mothers with Substance Use Disorder'. *Addictive Behaviors Reports*, 10: 100193. However, rather than a problem with the concept of certainty about other minds as contrary to mentalizing, this more likely reflects a methodological limitation: such certainties come in to play under conditions of high arousal, not as much when completing pen-and-paper questionnaires. Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Assessment of Mentalising', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 37–62, p. 57.

¹⁷⁸ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'from a research perspective the precise nature of the

Symptom contexts and networks

In diagnostic-centric practice within bureaucratically oriented healthcare systems or those structured by health insurance, symptoms may not be fully registered unless they can be understood as surface manifestations of an underlying diagnostic entity, with expectable causes, qualities, and prognosis. This is an analogy with physical illness, where a given symptom (e.g. coughing up blood) is understood to have a single cause, which will eliminate the symptoms when removed. This perspective is advocated by some mental health professionals and researchers. However, more structurally, it has been held in place in part by the wish of healthcare systems for an orderly and auditable response to individual needs and health trajectories.¹⁷⁹

For their part, Fonagy and colleagues have advocated in favour of a focus on symptoms and their interactions as the primary level of analysis.¹⁸⁰ This has led them to examine the role of the p-factor, as a transdiagnostic latent factor in the formation of mental health symptoms. They have, with varying degrees of caution, identified the p-factor with impairment in epistemic trust. As we saw epistemic trust, Fonagy and colleagues regard epistemic trust as a self-righting mechanism, which offers a means for learning, feedback, and adjustment in the face of challenges, as well as capitalization on available social supports—including the capacity to make effective use of mental health services. A reason for the occasional caution shown by Fonagy and colleagues in making this claim may well be that elsewhere they have considered various reasons why mental health symptoms might intersect and reinforce, which would contribute to a general latent factor. Epistemic trust may well be one contributor to the latent general dimension, but its identification with the p-factor would then be overhasty. Here we will identify six further factors in the writings of Fonagy and colleagues that might contribute to the existence of a latent factor of mental health symptoms.

A second reason why mental health symptoms may form a latent dimension is that the diverse conditions that prompt or maintain them may intersect and reinforce. This point has been a particular argument developed in Fonagy's writings with his wife Anna Higgitt in the 1990s and 2000s, reflecting her longstanding interest in health policy and work as a senior policy adviser to the Department for Health.¹⁸¹ Through research with prospective studies in the 1980s, Michael Rutter had first shown that most mental health conditions are best predicted not by particular kinds of risks but the aggregation of adversities or gaps in social support.¹⁸² These findings have subsequently been well replicated.¹⁸³ Fonagy and Higgitt

relationship between depression and BPD requires further elucidation ... there is uncertainty about the extent to which depression and BPD should be considered as part of the same spectrum' (p. 81).

¹⁷⁹ Rosenberg, C. E. (2002). 'The Tyranny of Diagnosis: Specific Entities and Individual Experience'. *The Milbank Quarterly*, 80(2): 237–260; Greco, M. (2016). 'What is the DSM? Diagnostic Manual, Cultural Icon, Political Battleground: An Overview with Suggestions for a Critical Research Agenda'. *Psychology & Sexuality*, 7(1): 6–22; Schnittker, J. (2017). *The Diagnostic System: Why the Classification of Psychiatric Disorders Is Necessary, Difficult, and Never Settled*, New York: Columbia University Press; Armstrong, D. (2019). 'Diagnosis: From Classification to Prediction'. *Social Science & Medicine*, 237: 112444.

¹⁸⁰ Cf. Borsboom, D. (2017). 'A Network Theory of Mental Disorders'. *World Psychiatry*, 16(1): 5–13.

¹⁸¹ Higgitt, A. (2000). 'Suicide Reduction: Policy Context'. *International Review of Psychiatry*, 12(1): 15–20.

¹⁸² Rutter, M. (1987). 'Psychosocial Resilience and Protective Mechanisms'. *American Journal of Orthopsychiatry*, 57(3): 316–331.

¹⁸³ See also McCrory, E. J. and Viding, E. (2015). 'The Theory of Latent Vulnerability: Reconceptualizing the Link between Childhood Maltreatment and Psychiatric Disorder'. *Development and Psychopathology*, 27(2): 493–505; Caspi, A., Houts, R. M., Belsky, D. W., Harrington, H., Hogan, S., Ramrakha, S., ... Moffitt, T. E. (2016). 'Childhood forecasting of a small segment of the population with large economic burden'. *Nature Human Behaviour*, 1: 5; Most recently, the 'Adverse Childhood Experiences' measure has served partly as a poster-child and partly as a placeholder for acknowledgement of the multiplier effect of risk in the course of development. See Steptoe, A.,

were struck by these findings, which stood in contrast to the psychoanalytic models of the day, in which early psychological risks were privileged as causes almost to the exclusion of the wider ecology of a person in the course of their development. They took from Rutter the conclusion that ‘risk conditions occur simultaneously’, and their ramification means that the accumulation of adversities or gaps in social support result not in an additive, but in a multiplying, likelihood of mental health symptoms of one form or another. The same was true of other outcomes such as criminal behaviour and substance use. The implication was that ‘outcomes for most etiologies exist along a continuum’.¹⁸⁴

A latent p-factor may find some of its basis, therefore, in the fact that the adversities that predispose or maintain mental illness in general can interest and reinforce one another. Fonagy and Higgitt offered the example of the intersection of social disadvantage and a very disappointing life event.¹⁸⁵ This intersection will create much more risk of various forms of mental illness than either of the risks alone, in part because the adversities may hit harder, and in part because each may knock out factors that would otherwise buffer the effects of the other. While individual qualities may offer protection to an extent, Fonagy and Higgitt claimed that this protection is always and necessarily lost in the face of accumulating risks: ‘we may expect children to be resilient to one or more of such stressors, but as risk compounds even the strongest constitutionally must succumb’. As adversities accumulate and supports decrease in the course of development, so too will various mental health symptoms. In recent work, Fonagy has highlighted that the association between ramifying developmental risk factors and later outcomes will have a diversity of mediators.¹⁸⁶ Some of the effects of adversity and lack of support on mental illness may be mediated by impairments in epistemic trust. Some of the effects, however, may be direct or operate through other mechanisms. As such, the coincidence of mental symptoms may in part reflect impairment of epistemic trust, but may also in part reflect the intersection and mutual reinforcement of adversities and gaps in social support.

A third reason that mental health symptoms may form a general latent factor can be extrapolated from the discussion of attachment and adaptation earlier in this chapter. Proposals by attachment researchers suggested to Fonagy and colleagues that human evolution may have pre-primed a repertoire of strategies and dispositions, which respond to escalating adversities and lack of support with coping strategies with the predictable outcome of short-term pay-offs but with a long-term price. Crittenden argued that humans have evolved to respond to dangerous and adverse circumstances with forms of perception and behaviour that exclude information about relationships that would hinder coping. In one class of response, individuals exclude emotional information prompted within relationships, because this information is regarded as disrupting the capacity to soldier on. In the other class of response, individuals exclude information about the temporal and causal sequencing of others’ availability, permitting the maintenance of vigilance. In both cases, however,

Marteau, T., Fonagy, P., and Abel, K. (2019). ‘ACEs: Evidence, Gaps, Evaluation and Future Priorities’. *Social Policy and Society*, 18(3): 415–424.

¹⁸⁴ Fonagy, P. and Higgitt, A. (2000). ‘Early Influences on Development and Social Inequalities: An Attachment Theory Perspective’, in A. R. Tarlov and R. F. S. Peter (eds), *The Society and Population Health Reader*, Volume 2: A State and Community Perspective, New York: New Press, pp. 104–130, p. 104.

¹⁸⁵ Fonagy, P. and Higgitt, A. (2007). ‘The Early Social and Emotional Determinants of Inequalities in Health’, in G. Baruch, P. Fonagy, and D. Robins (eds), *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, Chichester, UK: John Wiley & Sons, pp. 3–34, pp. 6–7.

¹⁸⁶ E.g. Steptoe, A., Marteau, T., Fonagy, P., and Abel, K. (2019). ‘ACEs: Evidence, Gaps, Evaluation and Future Priorities’. *Social Policy and Society*, 18(3): 415–424.

Crittenden theorized a spectrum of transformations of information about relationships, ranging through truth, error, omission, distortion, and deception. She held that diverse forms of mental health symptoms will increase, the further along this spectrum one goes. Reflecting on Crittenden's work, Fonagy reinterpreted this dimension as 'congruent' with a spectrum of epistemic mistrust.¹⁸⁷ Certainly, both models are centrally concerned with the extent to which information in relationships is processed accurately and experienced as relevant to the self. Nonetheless, it is not clear that all coping strategies accepting long-term costs for short-term pay-offs solely reflect epistemic trust.

Another attachment theorist who, as we saw, has thought about attachment in terms of strategies and long-term/short-term transactions is Belsky. As we saw adaptation and attachment, Belsky identified that insecure infant attachment is associated with early menarche and with more risk-taking behaviours. This offers support for his theory that evolution has primed humans to treat adverse early experience as a prime for biological and behavioural responses that prioritize short-term survival and reproduction, even if this may come with a long-term price. A portion of this price may be physiological, in terms of health implications. However, other aspects may register as mental health symptoms, such as anxiety, attentional problems, impulsivity, and conduct problems. The idea of strategies with expectable short-term benefits but long-term disadvantages may in some regards align with the concept of epistemic trust, in the proposal that human evolution has offered various ways that, even without the conscious intention of the individual, we may be disposed to respond to developmental adversity with less faith that all will work out. However, it does not seem likely that all the features of a life history strategy focused on short-term pay-off can be reduced to epistemic mistrust. The metabolic, social, and psychological strain and ramifications of this strategy may be regarded as another reason why symptoms may form a latent p-factor.¹⁸⁸

A fourth contributor to the p-factor as a latent dimension may stem from symptoms that arise to mitigate other symptoms, and so predictably co-vary. This was a classic argument of Winnicott's, who argued that some symptoms—such as depression—should be examined carefully to see whether they counterbalance other symptoms and result, overall, in greater psychological health than if they had been absent.¹⁸⁹ Though Fonagy and colleagues have not tended to follow Winnicott in this interpretation of depression, they have agreed on the general point. The example most often discussed by Luyten, Fonagy, and their collaborators is that overeating may function as a self-soothing or self-medicating strategy as a response to felt distress or emptiness. Prospective research has documented a sharp dose–response relationship between an individual's exposure to various adversities and their likelihood of disordered overeating.¹⁹⁰ This relationship may be underpinned in part by the role of serotonin and dopamine in both food and mood, but there is also clearly a wider context implicating

¹⁸⁷ Fonagy, P. (2016). 'The Role of Attachment, Epistemic Trust and Resilience in Personality Disorder: A Trans-Theoretical Reformulation'. DMM News, 26 September. Accessed at: <http://www.iasa-dmm.org/images/uploads/DMM%20%2322%20Sept%2016%20English.pdf>.

¹⁸⁸ The idea that sustained states of arousal may contribute to somatic symptoms without the involvement of epistemic mistrust is implied in Luyten, P., van Houdenhove, B., Lemma, A., Target, M., and Fonagy, P. (2012). 'A Mentalization-Based Approach to the Understanding and Treatment of Functional Somatic Disorders'. *Psychoanalytic Psychotherapy*, 26(2): 121–140, though no doubt epistemic mistrust may have a powerful role in sustaining these symptoms. Somatic symptoms load strongly with the p-factor in most factor analytic studies of symptoms.

¹⁸⁹ E.g. Winnicott, D. W. (1955). 'The Depressive Position in Normal Emotional Development'. *British Journal of Medical Psychology*, 28(2–3): 89–100.

¹⁹⁰ E.g. Kubzansky, L. D., Bordenois, P., Jun, H. J., Roberts, A. L., Cerda, M., Bluestone, N., and Koenen, K. C. (2014). 'The Weight of Traumatic Stress: A Prospective Study of Posttraumatic Stress Disorder Symptoms and Weight Status in Women'. *JAMA Psychiatry*, 71(1): 44–51.

attempts to regulate self–other relationships, feelings of privacy and self-ownership, as well as the cultural meanings of embodiment.¹⁹¹ Skårderud and Fonagy have argued that the symptoms of disordered eating should be regarded as stemming from a ‘need to drown out painful self-states’, and therefore ‘as attempts at recovering cohesion, vitality and self-regulation.’¹⁹² Equivalent arguments may be made for hypersomnia and insomnia, both of which may be initiated or sustained as a means to mitigate other mental health symptoms, and so will predictably co-occur with them.¹⁹³ Likewise, worry may develop as a response to other symptoms, as a preoccupation with attempting to resolve them. In fact, Fonagy and colleagues have identified that the absence of worry in the context of other mental health symptoms may come with its own problems, such as maintenance of risky behaviour without concern for possible consequences.¹⁹⁴ This aligns with the findings of anxiety as a positive prognostic sign in the Anna Freud Centre retrospective study (see Chapter 2).

A fifth contribution to the p-factor as a latent dimension may lie in reciprocally reinforcing relationships between mental health symptoms. The most salient of these for Fonagy and colleagues are the reciprocal relationships between problems in affect modulation, attentional difficulties, non-mentalizing, and a paucity of stable social relationships. In fact, Bateman and Fonagy observed that these are abstractions, and in reality each is likely a component of the others, with blurred lines between them. It is often the case that ‘one may end up triggering another,’ especially towards the more severe end of mental illness.¹⁹⁵ Reinforcing and intensifying interactions may in fact occur between symptoms that were initiated as attempts at coping. For instance over- or under-eating may begin as a coping strategy for mitigating rumination or feeling low, only to develop over time into a set of symptoms dominating the individual’s experience of food, or even of social interactions.¹⁹⁶ To take another example, reduced hours of sleep may have developed as a tactic to mitigate rumination or low mood that would otherwise occur during attempts to sleep. However, insufficient sleep is likely to contribute further over time to both symptoms. In fact, Patalay, Fonagy, and colleagues found that poor sleep was the single strongest loading on the p-factor of any mental health symptom, higher even than distress or anger.¹⁹⁷

Another set of reinforcing and intensifying relationships between symptoms have been discussed by Luyten, Fonagy, and colleagues in relation to severely depressed states. Severely depressed states may deplete the attentional focus required to clearly conceive of or reconsider the thoughts and feelings of oneself or others. This may prompt entry into a state of

¹⁹¹ Duschinsky, R., Reisz, S., and Messina, S. (2019). “Pulling the World in and Pushing it Away”: Participating Bodies and the Concept of Coping’. *Medical Humanities*, 45(2): 124–130.

¹⁹² Skårderud, F. and Fonagy, P. (2012). ‘Eating Disorders’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 347–384, p. 352.

¹⁹³ Blatt, S. J. and Luyten, P. (2009). ‘Depression as an Evolutionarily Conserved Mechanism to Terminate Separation Distress: Only Part of the Biopsychosocial Story?’. *Neuropsychoanalysis*, 11(1): 52–61, pp. 54–55.

¹⁹⁴ St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). ‘Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults’. *PloS One*, 12(4): e0175381.

¹⁹⁵ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 41.

¹⁹⁶ Luyten, P., Van Assche, L., Kadriu, F., Krans, J., Claes, L., and Fonagy, P. (2017). ‘Other Disorders often Associated with Psychological Trauma’, in C. Dalenber, S. Gold, and J. Cook (eds), *APA Handbook of Trauma Psychology*. Volume 1: Foundations in Knowledge, Washington, DC: American Psychological Association, pp. 261–262.

¹⁹⁷ Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., and Wolpert, M. (2015). ‘A General Psychopathology Factor in Early Adolescence’. *The British Journal of Psychiatry*, 207(1): 15–22: ‘Sleep disturbance emerges as a poor indicator of internalising problems and might be better conceived of as a generic indicator of vulnerability to psychiatric disorder’ (p. 19).

psychic equivalence, in which feeling low defines reality.¹⁹⁸ Or even when mental states are considered, in a state of affective hypoactivation use of reflections to account for and explain behaviour or experiences may be felt as futile.¹⁹⁹ In turn, modes of non-mentalizing will contribute to states of extreme certainty or extreme doubt about the meaning of relationships with other people or aspects of the self-representation relevant to action. This can directly prompt further feelings of depression, as well as responses such as social withdrawal that reinforce these symptoms and that hinder future mentalizing. Luyten, Fonagy, and colleagues have referred to such patterns as 'dysfunctional interpersonal transaction cycles'.²⁰⁰ However, they reported empirical findings that scores on the reflective function scale are not associated with depression scores among patients with severe depression. Instead, reflective function was negatively associated with the chronicity of depression and the likelihood of inpatient admission.²⁰¹ Such findings suggest that non-mentalizing may sustain depression, and mediate interpersonal transaction cycles, but is not itself its initial basis.²⁰² In the model proposed by Bateman and Fonagy, difficulties in forming and making use of relationships prompts feeling low, which in turn takes mentalizing offline and reinforces epistemic vigilance.²⁰³ This cascade hinders recovery from depression and increases the risk of relapse following recovery.

To take another example of disadvantageous interpersonal transaction cycles: Luyten, Fonagy, and colleagues have also proposed that substance and alcohol use may begin as a 'means of silencing painful thoughts'.²⁰⁴ However, in serving this function, and especially when drug and alcohol use are one of the few available tactics at the disposal of the individual for achieving short-term regulation, they may contribute over time to reciprocal relationships between problems with affect modulation, attentional difficulties, use of non-mentalizing modes, and a paucity of stable social relationships. St Clair and colleagues

¹⁹⁸ Fonagy, P. and Target, M. (2003). 'Being Mindful of Minds: A Homage to the Contributions of a Child-Analytic Genius'. *The Psychoanalytic Study of the Child*, 58(1): 307–321. 'Depression entails an over involvement with and concretisation of mood related ideation' (p. 319); Luyten, P., Lemma, A., and Target, M. (2019). 'Depression', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 387–401.

¹⁹⁹ Bateman and Fonagy describe how 'absence of arousal prevents the development of attachment-based affect, which is the area of sensitivity in interpersonal interactions for people with BPD. Treatment becomes cognitively organized and the patient is detached from relational process. Pretend mode is often associated with absence of affect and may become persistent. Working on areas of interpersonal sensitivity that lead to loss of mentalising becomes impossible because interpersonal meaning is absent.' Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 212.

²⁰⁰ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418, p. 407.

²⁰¹ Fischer-Kern, M., Fonagy, P., Kapusta, N. D., Luyten, P., Boss, S., Naderer, A., ... and Leithner, K. (2013). 'Mentalizing in Female Inpatients with Major Depressive Disorder'. *Journal of Nervous and Mental Disease*, 201(3): 202–207.

²⁰² See also Ensink, K., Bégin, M., Normandin, L., and Fonagy, P. (2016). 'Maternal and Child Reflective Functioning in the Context of Child Sexual Abuse: Pathways to Depression and Externalising Difficulties'. *European Journal of Psychotraumatology*, 7(1): 30611. Recent work has suggested qualifications to the Fischer-Kern study, in showing that distinct forms of depression may have different trajectories and correlates. Rost, F., Luyten, P., and Fonagy, P. (2018). 'The Anaclitic-Introjective Depression Assessment: Development and Preliminary Validity of an Observer-Rated Measure'. *Clinical Psychology & Psychotherapy*, 25(2): 195–209.

²⁰³ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 335–336.

²⁰⁴ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418, p. 400. See also Arefjord, N., Morken, K., and Lossius, K. (2019). 'Comorbid Substance Use Disorder and Personality Disorder', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 403–416.

examined associations between drug and alcohol use with the p-factor. Daily alcohol use was positively associated with the p-factor, but weakly ($r = .18$). Frequent alcohol consumption is normative in England, which may mask associations between the p-factor and drinking to silence painful thoughts (and perhaps feelings) in a brief self-report measure. Or painful thoughts and feelings may prompt alcohol use in the general population without contributing to the p-factor.

By contrast, the association between the p-factor and various forms of substance use was much more substantial ($r = .34$ for cannabis and $.37$ for other illegal drugs).²⁰⁵ Skårderud and Fonagy have described the ‘polysemy’ of behaviour, its embeddedness in multiple systems of causation and potential meaning-making. They used this idea to argue against interpretations of particular mental health symptoms as having any ‘one’ meaning.²⁰⁶ In this light, there may be such polysemy in why substance use is associated with other mental health problems:

- Epistemic mistrust may contribute to a sense of risks as irrelevant, credulity towards peer norms, difficulties learning lessons for future behaviour from experiences with negative consequences, and difficulties in seeking or making use of social or medical support for substance addiction.
- The multiplier effect of risk may contribute to the availability of substances, the presence of substance-using peers, and contexts in which drug use is frequent.
- Life histories primed for short-term benefits may contribute to substance use through impulsivity, feelings of urgency in response to distress, and intensified neuroendocrinal rewards.
- Substance use may offer significant short-term benefits in self-medicating for other symptoms such as distress or rumination
- Substance use may also contribute to the cause or maintenance of other mental health problems, as well as reducing the effectiveness of protective factors—for instance, by hindering mentalising.

As the illustration of substance use suggests, there may be a variety of processes that contribute to the existence of a general latent dimension for mental health symptoms. The claim by Fonagy and Campbell that the ‘the p factor is a measurement of epistemic trust’²⁰⁷ is over-strong. Evidence in favour of a link between the two is still emerging, and there seem to be a variety of other processes in play. However, the considerations elsewhere in the work of these researchers and their collaborators offer the basis for an account of symptom contexts and networks in clustering mental health symptoms together into a latent dimension. The ideas present in the work of Fonagy and colleagues includes insights from cognitive and

²⁰⁵ St Clair, M. C., Neufeld, S., Jones, P. B., Fonagy, P., Bullmore, E. T., Dolan, R. J., ... and Goodyer, I. M. (2017). ‘Characterising the Latent Structure and Organisation of Self-Reported Thoughts, Feelings and Behaviours in Adolescents and Young Adults’. *PloS One*, 12(4): e0175381. The strength and direction of this association may well be influenced by cultural factors, though see also Handeland, T. B., Kristiansen, V. R., Lau, B., Håkansson, U., and Øie, M. G. (2019). ‘High Degree of Uncertain Reflective Functioning in Mothers with Substance Use Disorder’. *Addictive Behaviors Reports*, 10: 100193.

²⁰⁶ Skårderud, F. and Fonagy, P. (2012). ‘Eating Disorders’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 347–384, p. 356.

²⁰⁷ Fonagy, P. and Campbell, C. (2015). ‘Bad Blood Revisited: Attachment and Psychoanalysis, 2015’. *British Journal of Psychotherapy*, 31(2): 229–250, p. 243.

information processing (epistemic mistrust), developmental psychopathology (multiplier effect of risk), evolutionary socio-biology (life history theory), strengths-based approaches (self-medication) as well as network approaches (reciprocal instigation and reinforcement). All these processes may also be relevant to how amenable symptoms are to mental health intervention. The work by Fonagy and colleagues to apply their ideas to therapeutic practice will be the focus on the next chapter.

Mentalization-based therapies

Introduction

From early in his career, Fonagy was adamant that ‘it is inconceivable’ that ‘if the ideas proposed by the group have substance’, that they would have no implications for the technical priorities of therapeutic work with patients.¹ As we saw in Chapter 1, Fonagy’s clinical work in the 1980s with patients diagnosed with borderline personality disorder (BPD) played an important role in the development of his thinking. Clinical work as a psychoanalyst working with adults and children directly informed his concern with mentalizing, which grew out of technical innovations in ‘developmental help’ offered at the Anna Freud Centre. In turn, the theory developed by Fonagy and colleagues fed into the emergence of mentalization-based therapy (MBT) as a treatment modality. Initially, in the 1990s, the target of this intervention was patients with BPD, a group who were often regarded as ‘untreatable’. However, MBT has subsequently been used with patients facing a variety of mental health symptoms, including common symptoms like anxiety and depression. In this chapter, we will describe the development of MBT over time, including variants such as dynamic interpersonal therapy. We will also consider some remaining questions about MBT, including the potential for MBT to harm particular groups of patients.

The development of mentalization-based therapy

Fonagy has recalled the origin of MBT in clinical discussions with George Moran in the late 1980s. Moran was the charismatic director of the Anna Freud Centre, and had a practice with adolescents with poor control of their diabetes. Fonagy and Moran used to meet on Saturdays to discuss and seek advice on their adolescent clinical cases: ‘We gradually realized that all the interpretations we were giving didn’t make any difference to these young people. But once we started talking to them about their lives and what their thoughts and feelings were in very simple language—their minds changed and their diabetic control got better.’² Moran’s approach aligned with aspects of Hurry’s own clinical technique in her analysis of

¹ Fonagy, P. (1998). ‘Moments of Change in Psychoanalytic Theory: Discussion of a New Theory of Psychic Change’. *Infant Mental Health Journal*, 19(3): 346–353, pp. 351–352.

² Ezrati, O. (2014). ‘Freud Off: Giving New Meaning to Psychoanalysis’, *Haaretz*, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>. Another retrospective account is offered by Fonagy and colleagues in their REF 2014 Impact Case Study: ‘The research originated in the 1990s when Fonagy and colleagues discovered that juvenile diabetics with poor insulin control struggled to depict their social experiences accurately in mental state terms and that this lack of “mentalising” capacity meant they often could not predict the consequences of their own and others’ actions. Listening to them talk about their feelings and their understanding of feelings dramatically improved their diabetic control (measured through glycosylated haemoglobin levels). Clinical work both with these young people and BPD patients suggested that failure of mentalising often followed a combination of early neglect and childhood trauma.’ University College London (2014). ‘REF 2014: Psychology, Psychiatry and Neuroscience Impact Case Study’. Accessed at: [https://results.ref.ac.uk/\(S\(jj2mvb3fbee3zpqblartx2d\)\)/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202](https://results.ref.ac.uk/(S(jj2mvb3fbee3zpqblartx2d))/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202).

Fonagy as an adolescent, and with Fonagy's thinking about adult patients such as Mr S. (see Chapter 1). These experiences revealed the potential benefits, including relief of symptoms, that seemed to stem from patients' gaining perspective on the sources and scope of their feelings and thoughts.³

In the early 1990s, Fonagy had become convinced of the value of manualized approaches to therapeutic intervention: for the sake of conducting research evaluation; to facilitate transparency to stakeholders and commissioners of services; in order to establish an explicit model for teaching and further improvement; and to help buffer the effects of clinicians' own anxieties and the interpersonal pressures generated in the consulting room.⁴ He perceived a gap in available treatment modalities for BPD. The condition was regarded by many as untreatable.⁵ And cognitive behavioural therapy (CBT), otherwise the dominant approach in UK mental health services, was struggling to demonstrate effectiveness with this group of patients.⁶ Bateman and Fonagy regarded mentalization as the ingredient within the psychoanalytic approach central to its treatment efficacy. Psychoanalytic therapies are especially concerned with exploring thoughts and feelings, comparing experiences in the present with constructions about the past, and reducing the potential for overwhelming experiences that might put at risk the capacity to reflect—for instance, through use of the couch.⁷ However, the mentalization-focused elements of psychoanalytic therapies could be refined, and also extracted from those elements that distract from or hinder mentalizing. As discussed in previous chapters, these latter include the use of speculative interpretations, use of sexuality as a metaphor for other challenges and conflicts, and clinicians' opacity and reticence with the patient.

In a study published in 1999, Bateman and Fonagy compared the outcomes for 19 patients with BPD treated with an early version of MBT, and 19 matched controls who received treatment as usual under the NHS. The key elements of the treatment were weekly individual MBT and thrice-weekly group analytic psychotherapy. At both 6 months and 18 months after the end of treatment, the patients who received the new therapy had fewer depressive symptoms, engaged in fewer suicidal and self-harming acts, spent fewer days as inpatients, and reported better social and interpersonal functioning.⁸ In fact, the patients who had received MBT continued to improve over time, whereas the controls did not.⁹ Though the early version of MBT was initially more expensive, this money was ultimately recouped by follow-up

³ For recent evidence supporting Moran's findings with patients with brittle diabetes, see Costa-Cordella, S., Luyten, P., Cohen, D., Mena, F., and Fonagy, P. (2020). 'Mentalizing in Mothers and Children with Type 1 diabetes'. *Development and Psychopathology*, Early View.

⁴ Fonagy, P. and Target, M. (1996). 'Should we Allow Psychotherapy Research to Determine Clinical Practice? Comments on Sol J. Garfield: "Some Problems Associated with "Validated" Forms Of Psychotherapy"'. *Clinical Psychology: Science and Practice*, 3: 245–250; Fonagy, P. (1999). 'Achieving Evidence-Based Psychotherapy Practice: A Psychodynamic Perspective on the General Acceptance of Treatment Manuals'. *Clinical Psychology: Science and Practice*, 6(4): 442–444.

⁵ In some quarters, these attitudes towards personality disorder have continued. See e.g. Chartonas, D., Kyratsous, M., Dracass, S., Lee, T., and Bhui, K. (2017). 'Personality Disorder: Still the Patients Psychiatrists Dislike?'. *BJPsych Bulletin*, 41(1): 12–17; Beryl, R. and Völlm, B. (2018). 'Attitudes to Personality Disorder of Staff Working in High-Security and Medium-Security Hospitals'. *Personality and Mental Health*, 12(1): 25–37.

⁶ Leichsenring, F. and Leibing, E. (2003). 'The Effectiveness of Psychodynamic Therapy and Cognitive Behavior Therapy in the Treatment of Personality Disorders: A Meta-Analysis'. *American Journal of Psychiatry*, 160: 1223–1232.

⁷ For further discussion of the mentalizing elements of psychoanalysis, see especially Fonagy, P. and Adshad, G. (2012). 'How Mentalisation Changes the Mind'. *Advances in Psychiatric Treatment*, 18(5): 353–362.

⁸ Bateman, A. and Fonagy, P. (1999). 'Effectiveness of Partial Hospitalization in the Treatment of Borderline Personality Disorder—A Randomized Controlled Trial'. *American Journal of Psychiatry*, 156: 1563–1569.

⁹ Bateman, A. and Fonagy, P. (2001). 'Treatment of Borderline Personality Disorder with Psychoanalytically Oriented Partial Hospitalization: An 18-Month Follow-Up'. *American Journal of Psychiatry*, 158: 36–42.

in terms of a large reduction in the patients' use of community and hospital services.¹⁰ Even eight years after the initial intervention, the researchers found that the patients who received MBT showed fewer symptoms, less suicidality, and were more likely to be in employment or education. They had less crisis service use, less use of long-term outpatient mental health services, and were less likely to have social services involvement.¹¹ Many patients who received MBT still faced substantial difficulties requiring forms of professional support, but this was at a lower frequency and over shorter duration than for control participants.

In the early 2000s, Bateman and Fonagy developed a manualized MBT approach, published as a book in 2004.¹² The process by which this manual was developed is relatively opaque.¹³ However, in texts from the period, Bateman and Fonagy reflected that 'the overall goals of treatment are to stabilize the self-structure through the development of stable internal representations, formation of a coherent sense of self, and capacity to form secure relationships. But the self-structure is destabilized in the context of emotional turmoil and so a further goal is identification and appropriate expression of affect.'¹⁴ These goals were sought in the new modality by carefully supporting patients to articulate and consider their thoughts and feelings and those of others, without becoming overwhelmed, impatient, or entering into pretend mode.

In general terms, psychoanalytic approaches to therapy often centre attention on unconscious conflicts and fantasies, and the role of defences in holding disturbing thoughts or feelings out of consciousness. By contrast, MBT prioritizes present-day experience and near-conscious thoughts and feelings that can, perhaps with help, be identified by the patient themselves and considered together with other experiences.¹⁵ The training requirements for an MBT therapist are very modest compared with psychoanalytic training, contributing to greater scalability.¹⁶ And whereas classical psychoanalytic technique encouraged reticence,

¹⁰ Bateman, A. and Fonagy, P. (2003). 'Health Service Utilization Costs for Borderline Personality Disorder Patients Treated with Psychoanalytically Oriented Partial Hospitalization versus General Psychiatric Care'. *American Journal of Psychiatry*, 160(1): 169–171.

¹¹ Bateman, A. and Fonagy, P. (2008). '8-Year Follow-Up of Patients Treated for Borderline Personality Disorder: Mentalization-Based Treatment Versus Treatment As Usual'. *American Journal of Psychiatry*, 165(5): 631–638; Bateman, A., Constantinou, M. P., Fonagy, P., and Holzer, S. (2020). 'Eight-Year Prospective Follow-Up of Mentalization-Based Treatment versus Structured Clinical Management for People with Borderline Personality Disorder'. *Personality Disorders: Theory, Research, and Treatment*, Early View. A first empirical study of the role of mentalization within treatment was conducted by Chiesa and Fonagy, who studied 300 patients in personality disorder services: from the Cassel Hospital Residential inpatient programme; from a therapeutic outpatient programme; and management as usual by the Devon personality disorder services. None of these services were mentalization-based treatments. However, Chiesa and Fonagy found that patient reflective functioning on the Adult Attachment Interview was improved by the outpatient programme, which also had the best patient outcomes by a six-year follow-up. By this point, 62% of patients in the outpatients intervention were below the clinical cut-point for a personality disorder, compared with 26% for the inpatient intervention and 13% in treatment as usual. Chiesa, M., Fonagy, P., and Holmes, J. (2006). 'Six-Year Follow-Up of Three Treatment Programs to Personality Disorder'. *Journal of Personality Disorders*, 20(5): 493–509.

¹² Bateman, A. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorder: Mentalization-Based Treatment*, Oxford: Oxford University Press.

¹³ Some historical remarks are presented in Fonagy, P. and Bateman, A. (2009). 'A Brief History of Mentalisation-Based Treatment and its Roots in Psychoanalytic Theory and Practice', in M. Brownescombe Heller and S. Pollet, (eds), *The Work of Psychoanalysts in the Public Health Sector*, London: Routledge, pp. 156–176.

¹⁴ Bateman, A. W. and Fonagy, P. (2003). 'The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder'. *Bulletin of the Menninger Clinic*, 67: 187–211, p. 195.

¹⁵ Sandler can be regarded as an intermediary figure here, because his account has elements that correspond much more to a classic psychoanalytic account (e.g. drives, discussion of the ego), and some elements that seem to presage MBT, such as his central concern with the preconscious. Fonagy, P. (2005). 'An Overview of Joseph Sandler's Key Contributions to Theoretical and Clinical Psychoanalysis'. *Psychoanalytic Inquiry*, 25(2): 120–147.

¹⁶ This may be one contributing reason why the British Psychoanalytic Council have treated Fonagy's research and teaching on MBT as invalid activity for the purposes of continuing professional development as a psychoanalyst: despite many points of commonality, MBT is not framed as operating on the same epistemic object as

MBT adopted a more collaborative relationship between therapist and patient, with greater attempt by the therapist to make his or her thoughts and feelings legible for discussion and consideration. This collaboration is initiated through active pursuit of the inquisitive stance, expressions of curiosity, and the elicitation of detailed descriptive accounts by the patient of their experiences. The tone of interactions should aim to support the patient in achieving some degree of emotional equilibrium—sufficient to find the therapeutic encounter safe enough to consider thoughts and feelings.

Bateman and Fonagy argued that consideration of thoughts and feelings should place them as fully as possible in context, in order to understand their applicability in accounting for the patient's behaviour and experience, as well as the behaviour and experience of others. Together, the patient and therapist attempt to elaborate upon descriptions, identifying causal sequences and the role played by thoughts and feelings within them. Within this process, 'simply looking at a feeling and its antecedents and consequences is not enough. The patient must be helped to consider who engendered the feeling and how, to explore whether the feelings have occurred or are connected to events either in the recent or longer-term past, to assess the appropriateness of the feeling to any given situation in terms of others' understanding of the patient, and to establish the appropriate locus of these feelings within current relationships, either past or present.'¹⁷ As part of this process, patients are asked to compare their thoughts and feelings at different times, as a step towards recognizing their causes and consequences.

In the course of therapy, the patient should be supported to articulate and consider mental states—even when, and especially when, intense emotions are aroused. This is part of the specific importance attributed to group-based aspects of MBT, because group conditions arouse strong emotions for patients and provide a supported opportunity to attend to their own and others' mental states. Likewise, the relationship with the therapist is also used as a prompt for patients to identify and examine their thoughts and feelings, especially in the context of considering their assumptions about other people and characteristic patterns of relating to them. Patients' experiences of the group and of the therapist should first be validated, then explored and elaborated further in collaboration. Alternative or additional accounts can then be considered, as well as the patient's reaction to these perspectives.

Dynamic Interpersonal Therapy

In the late 2000s, opportunities arose for embedding psychotherapies within primary care in the UK, thanks to the Improving Access to Psychological Therapies initiative.¹⁸ This led to a large expansion of the availability of low-intensity CBT. Several psychodynamically inspired approaches also sought to be adopted within UK primary care. This included cognitive analytic therapy (CAT),¹⁹ and the Tavistock model.²⁰ Fonagy and colleagues saw an opportunity

psychoanalytic approaches. Fonagy, P. and Allison, E. (2016). 'Commentary on Kernberg and Michels'. *Journal of the American Psychoanalytic Association*, 64(3): 495–500.

¹⁷ Bateman, A. W. and Fonagy, P. (2003). 'The Development of an Attachment-Based Treatment Program for Borderline Personality Disorder'. *Bulletin of the Menninger Clinic*, 67: 187–211, pp. 196–197.

¹⁸ Clark, D. M. (2018). 'Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program'. *Annual Review of Clinical Psychology*, 14.

¹⁹ Ryle, A. and Kerr, I. B. (2003). *Introducing Cognitive Analytic Therapy: Principles and Practice*, John Wiley & Sons.

²⁰ Malan, D. and Della Selva, P. C. (2007). *Lives Transformed: A Revolutionary Method of Dynamic Psychotherapy*, Karnac Books.

for MBT, if it could be successfully adapted for delivery in primary care. In 2011, Lemma, Target, and Fonagy published the manual for a brief form of MBT, focused specifically on depression and on 'anxiety with low mood'.²¹ They termed this approach 'dynamic interpersonal therapy' (DIT). In the model of clinical action underpinning DIT, depression and 'anxiety with low mood' are prompted by problems in current close relationships.²² In particular, experiences of impending or actual separation, rejection, loss, or failure were identified as particularly important threats to close relationships if they hinder the capacity of an individual to use those relationships as a secure base and safe haven. The interaction of relationship problems and negative affect may also prompt use of non-mentalizing modes—though generally not to the same degree or pervasiveness as in personality disorders:

DIT's starting point is rooted in the common clinical observation that patients who present as depressed and/or anxious almost invariably also present with difficulties and distress about their relationships. Although the patient may well experience his problem as 'I cannot sleep or concentrate' or 'I can't face going into crowded places, or going to work', the DIT therapist reframes such symptoms of anxiety and depressions as manifestations of a relational disturbance, which the patient cannot understand, or understands in a maladaptive way, attributing to himself or others motivations which are unlikely or unhelpful. Once the patient is helped to make some changes in the way he approaches relationship difficulties, depressive and anxious symptoms are typically alleviated.²³

DIT does not attempt to improve all aspects of mentalizing. The focus is on the patient's capacity to conceive of and reconsider the thoughts and feelings implicated in their present-day perceptual experience and social behaviour. As a brief treatment, the intervention focuses on improving 'the patient's capacity to reflect on his own states of mind', and the patient's understanding of 'the connection between his presenting symptoms and what is happening in his relationships'.²⁴ This is achieved over five steps. In the first step, the patient and the therapist identify a problem in one or more of the patient's relationships, perceived as contributing to depression and/or anxiety. By the end of the early sessions of therapy, the therapist and patient should have at least a tentative understanding of the kinds of relationship the patient tends to create and how this relates to their presenting symptoms. In the second step, the therapist works collaboratively with the patient to create a picture of the thoughts and feelings raised by the problem. In the third step, the therapist encourages the patient to explore alternative ways of considering these thoughts and feelings. The patient-therapist relationship is utilized to help the patient identify their characteristic assumptions about relationships, and ways of responding to others. Fourth, the therapist supports the patient to reconsider these assumptions, and their associated thoughts and feelings. Finally, the

²¹ Lemma, A., Target, M., and Fonagy, P. (2011). *Brief Dynamic Interpersonal Therapy: A Clinician's Guide*, Oxford: Oxford University Press.

²² See also Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418: 'The basic assumption of the mentalisation-based approach to depression is that depressive symptoms reflect responses to threats to attachment relations, and thus threats to the self, either because of (impending) separation, rejection or loss; (impending) failure experiences; or a combination of both ... Moreover depressed mood leads to further increases in arousal and stress levels, resulting in further impairments and distortions in mentalisation, which in turn lead to a loss of resilience' (p. 389).

²³ Lemma, A., Target, M., and Fonagy, P. (2011). *Brief Dynamic Interpersonal Therapy: A Clinician's Guide*, Oxford: Oxford University Press, p. 52.

²⁴ *Ibid.* 63.

therapist presents the patient with a written summary of the collaboratively held view of the shift in perspective achieved over the previous sessions.

Lemma, Target, and Fonagy gave self-representations (see Chapter 6) a fundamental place in DIT. They proposed that short-term work with patients should be oriented by an ‘interpersonal and affective focus.’²⁵ This focus comprises some representation about the self, some representation about others, the emotion that links them, and the psychological function of the configuration. So, for instance, a relevant self-representation may be the feeling that ‘I always ask for too much.’ The representation of others may be ‘They are not there when I need them.’ The linking emotion may be worry about abandonment. And the psychological function of the configuration may be a pre-emptive rejection of and distrust in others, which protects the individual from hurt and also expresses frustration with others, while also itself contributing to the frustration-evoking situation.

The DIT therapist should not take as their direct aim to improve the accuracy, coherence, or consistency of the patient’s self-representation—for instance, by helping the patient calibrate how much exactly they should ask for from others. Instead, the therapist should begin by asking the patient to describe and characterize the representation of self and other in detail, to produce a shared vocabulary for talking about these.²⁶ For Lemma, Target, and Fonagy, the ‘self’ of the patient is not treated as the primary target of therapeutic intervention. Nonetheless, representations of the self are considered for the contribution they make to characteristic forms of relating, the individual’s perceptual experience is taken as a key object for reflection, and support for mentalizing seeks to sustain the patient’s capacity for attention to their own thoughts, feelings, and/or intentions.

DIT is in several regards closer to cognitive behavioural therapy or cognitive analytic therapy than to MBT.²⁷ This reflects its development for use in primary care mental health delivery and patient groups. Four particular contrasts can be drawn between MBT and DIT. In DIT, the focus of the work is localized to depression or anxiety with low mood; therapy is always provided individually rather than sometimes using a combination of individual and group work; the focus is on mentalizing the self and there is, in principle, little explicit focus on supporting the patient to mentalize the thoughts and feelings of others; and the therapeutic work is less concerned with managing the potential for outbursts and lability of sadness and anger, or helping the patient mentalize in the context of distress. Nonetheless, Fonagy has argued that DIT still bears the defining marks of a mentalization-based therapy:²⁸

- 1) Maintain and, when it is lost, regain mentalizing (in both parties).
- 2) Active, curious, inquisitive; don’t feign understanding.
- 3) Direct joint attention to mental states.
- 4) Ordinary/non-expert: avoid guise of privileged knowledge about patients’ mind.

²⁵ Ibid. 107–108.

²⁶ Ibid. 116.

²⁷ Bateman and Fonagy argue that CAT, unlike MBT, lacks a developmental perspective and does not use psychoanalytic interpretations. These criticisms seem arguable as grounds of contrast, both for MBT and DIT. Bateman, A. W. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorders: Mentalization Based Treatment*, Oxford: Oxford University Press, pp. 129–132. In recent years, there has been some mention of cognitive mentalising therapy, as an explicit attempt to blend MBT and CBT. However, references have generally been passing, and there has been no justification offered regarding the intended benefits of cognitive mentalising therapy compared with its parent modalities. The fullest description to date appears in Bales, D. L. and Bateman, A. W. (2012). ‘Partial Hospitalization Settings’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 197–227.

²⁸ Fonagy, P. (2019). ‘Antisocial PD and Mentalisation Based Treatment.’ Paper presented on 3 July to the International Congress of the Royal College of Psychiatrists, London.

- 5) Emphasis on perspective-taking and marking discrepancies between perspectives (and exploring their sources).
- 6) 'Not knowing' stance: eschew certainty, mark what is not obvious but is presented that way, mark when you do suspect you know.
- 7) Model active, intentional effort to find out about opaque mental life.
- 8) Humility—acknowledge one's own non-mentalizing errors, model interest in being corrected and having one's mind changed.
- 9) Doggedness around exploring misunderstandings.
- 10) Transparency and authenticity about confusion, puzzlement, and self-reflection.

Many of these are generic aspects of relational psychotherapy, as Fonagy and colleagues have readily acknowledged. To the extent that there is a distinctively MBT quality to DIT, it perhaps can be identified in the proposed focus on sustaining and modelling generative doubt and the inquisitive stance, and avoidance of certainties that can inhibit these. This includes certainties prompted by holding one perspective to the exclusion of others. It also includes certainties prompted by the status of the therapist, and the assumption that he or she has privileged knowledge. The clinician's mental processes, in their potential wildness, are not made wholly transparent to the patient. However, the clinician does seek to articulate experiences of generative doubt and reconsideration so that the patient can see how this works and why it may be valued.

Though CBT remains the dominant approach, DIT has become a widely adopted treatment modality within UK primary care, as have some other low-intensity psychodynamic therapies such as CAT. Whether the focus on sustaining and modelling generative doubt, which could be argued to distinguish DIT at a theoretical level, leads to actual differences in clinical practice between DIT and other low-intensity psychodynamic therapies remains unknown. Until clear evidence to the contrary is published, we suspect that there is reason for gentle scepticism on this front.

Mentalization-Based Therapy today

Over time, MBT has become a widely used treatment modality in the treatment of personality disorders, with DIT capturing a further share of the market for primary care mental health services.²⁹ By 2014, over two and a half thousand clinicians had been trained in a form of MBT.³⁰ Versions of MBT have been developed and delivered for patients referred on the basis of various diagnoses, including eating disorders, depression, conduct disorder, and anti-social personality disorder. MBT has also been adapted for delivery with children, and as a form of family therapy.³¹ MBT programmes have become institutionally embedded in

²⁹ Work is also underway to elaborate and test a longer version of DIT for chronic depression with attendant other mental health symptoms, especially relevant to cases where other treatments have been attempted without success: the researchers consider the 'treatment-resistant features of the complex case of depression to be rooted in difficulties with mentalizing and epistemic mistrust', Rao, A. S., Lemma, A., Fonagy, P., Sosnowska, M., Constantinou, M. P., Fijak-Koch, M., and Gelberg, G. (2019). 'Development of Dynamic Interpersonal Therapy in Complex Care (DITCC): A Pilot Study'. *Psychoanalytic Psychotherapy*, 33(2): 77–98, p. 82.

³⁰ University College London (2014). 'REF 2014: Psychology, Psychiatry and Neuroscience Impact Case Study'. Accessed at: [https://results.ref.ac.uk/\(S\(jj2mvvb3fbee3zpqblartx2d\)\)/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202](https://results.ref.ac.uk/(S(jj2mvvb3fbee3zpqblartx2d))/DownloadFile/ImpactCaseStudy/pdf?caseStudyId=44202).

³¹ Midgley, N., Ensink, K., Lindqvist, K., Malberg, N., and Muller, N. (2017). 'Mentalization-Based Treatment for Children: A Time-Limited Approach'. *American Psychological Association*. Accessed at: <https://doi.org/10.1037/0000028-000>; Asen, E. and Midgley, N. (2019). 'Mentalization-Based Approaches to Working with

many services. For instance, an MBT programme for antisocial personality disorder (ASPD) has been rolled out across the whole of the Irish prison service.³² Even before the COVID-19 pandemic initial attempts had been made to offer MBT via phone or video-link, and in general the modality has proved relatively straightforward to adapt for remote delivery. One challenge has been that Fonagy and colleagues feel that the pandemic has spurred greater epistemic hypervigilance, making it more difficult for many patients to trust and learn from encounters with the clinician. Another challenge has been that many of the therapist's non-verbal ostensive cues are hindered or blocked by remote communication. However these challenges have been met through an intensified focus on what ostensive cues remain available for signalling recognition of mental states and modulating the patient's level of arousal.³³ Writing in 2016, Bateman and Fonagy remarked that 'MBT has been more successful than we ever anticipated'.³⁴ They identified several reasons for this growing popularity over the past two decades.

One has been that the modality had a research base. The 1999 study by Bateman and Fonagy, showing clinical benefits and cost-effectiveness despite its small sample, has proven an especially potent resource for the approach over subsequent decades in arguing for use of the modality in cash-strapped and hard-pressed services. Later studies have shown aligned findings, even if reported effects have been weaker.³⁵ If CBT certainly remains dominant within contemporary mental health services, its credentials as an evidence-based treatment have helped MBT find several relevant niches. For instance, 'clinicians find that some families are not able to benefit fully from the strategies they are offered because they do not

Families', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 136–149.

³² Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Antisocial Personality Disorder in Community and Prison Settings', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 335–349, p. 347.

³³ Fonagy, P., Campbell, C., Truscott, A., and Fuggle, P. (2020). 'Debate: Mentalising remotely—The AFNCCF's adaptations to the coronavirus crisis'. *Child and Adolescent Mental Health*, 25(3): 178–179. Ventura Wurman, T., Lee, T., Bateman, A., Fonagy, P., and Nolte, T. (2020). 'Clinical management of common presentations of patients diagnosed with BPD during the COVID-19 pandemic: the contribution of the MBT framework'. *Counselling Psychology Quarterly*, Early View. Fisher, S., Guralnik, T., Fonagy, P., and Zilcha-Mano, S. (2020). 'Let's face it: video conferencing psychotherapy requires the extensive use of ostensive cues'. *Counselling Psychology Quarterly*, Early View. Lassri, D. and Desatnik, A. (2020). 'Losing and regaining reflective functioning in the times of COVID-19: Clinical risks and opportunities from a mentalizing approach'. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1): S38.

³⁴ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. v.

³⁵ The findings from the 1999 study stand out in part against the relatively poor backdrop of services at the time. Later trials of MBT have found weaker effects, as have all other trials of treatments for personality disorders. Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., and Cuijpers, P. (2017). 'Efficacy of Psychotherapies for Borderline Personality Disorder: A Systematic Review and Meta-Analysis'. *JAMA Psychiatry*, 74(4): 319–328. The most probable interpretation is that treatment as usual has converged with the components of many of the branded therapies, making it more difficult for branded interventions to demonstrate effectiveness, though other factors may certainly also be implicated in these declines. Fonagy has also speculated that diagnosis-focused and manualized therapies may have also hindered efficacy by reducing attention to the specificities of the patient and their mental states, perhaps even encouraging entry into teleological mode. Publication bias and selective reporting may also play a role. For discussions, see Young, N. S., Ioannidis, J. P., and Al-Ubaydli, O. (2008). 'Why Current Publication Practices may Distort Science'. *PLoS Medicine*, 5(10): e201; Fonagy, P., Luyten, P., and Bateman, A. (2017). 'Treating Borderline Personality Disorder with Psychotherapy: Where Do We Go From Here?'. *JAMA Psychiatry*, 74(4): 316–317; Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast/>; Weisz, J. R., Kuppens, S., Ng, M. Y., Vaughn-Coaxum, R. A., Ugueto, A. M., Eckshtain, D., and Corteselli, K. A. (2019). 'Are Psychotherapies for Young People Growing Stronger? Tracking Trends over Time for Youth Anxiety, Depression, Attention-Deficit/Hyperactivity Disorder, and Conduct Problems'. *Perspectives on Psychological Science*, 14(2): 216–237.

have, or struggle to maintain, the capacity for reflective function and emotional regulation that are a prerequisite for their use.³⁶ In such contexts, MBT is attractive as an alternative evidence-based modality. A second factor identified by Bateman and Fonagy as contributing to the popularity of MBT is that MBT overlaps considerably with other treatment modalities and does not require a high level of specialism to be delivered in an adherent manner. It requires hard work, skill, and practice for a therapist to deftly check their understanding of the patient's mental state, and whether this corresponds with the patient's understanding. But it is an elaboration of the skills required by ordinary life, and a more sharply developed version of skills common to most therapeutic modalities.

A further reason Bateman and Fonagy provided for the popularity of MBT is that 'clinicians easily understand the ideas underpinning the model and recognise that promoting mentalising is something they are already doing in their clinical work'. Clinicians appear to experience MBT as a 'map through the woods', helping clinical problems make sense and providing guidance for how to approach them. Though, at the same time, Bateman and Fonagy admit that 'over the past few years it has become apparent that we were not specific enough about some of the core components of the model'.³⁷ True, as Sandler observed (see Introduction), blurry concepts can be helpful for handling blurry phenomena. Nonetheless, there may also be a price in terms of coherence, and articulation of causal processes. Indeed, part of the appeal of 'mentalizing' may have stemmed from the way that the concept can magnetize various investments and meanings, so that the interpretations of different clinicians have at most a family resemblance to one another. One clinician may have supported patients to conceive of feelings implicated in the motivations and intentions of others; another may have supported patients to reconsider their own thoughts in making sense of their past experiences; another may have supported patients to gain skills in reconsidering the observable social behaviour of others and its possible meanings. Any of these, and many others, could legitimately say that they were implementing one of the definitions of mentalizing offered by Fonagy and colleagues (see Chapter 4). And any of these, and many others, would be judged to be showing fidelity to the modality if assessed by Fonagy's 10 defining markers. Similarly, these different forms of clinical practice would be judged as adherent on the MBT Adherence and Competence Scale.³⁸

A final contribution to the popularity of MBT, listed by Bateman and Fonagy, is that the approach has broad possible applications because it integrates both developmental psychology and social cognition. Again, though, it should be noted that broad applicability has to an extent been purchased at the expense of—mostly inadvertent—construct variance across domains. As we saw in the previous chapter, for example, Fonagy's 2018 model integrated developmental psychology and social cognition, but in doing so included the concept of epistemic trust twice, and with different meanings on each occasion. So, for instance, one clinician may encourage young children to engage in symbolic play and then help the patient identify feelings implicit in these narratives; another clinician may help adolescents re-evaluate their sense of others' thoughts about them; a third clinician may focus on helping

³⁶ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

³⁷ Ibid.

³⁸ Karterud, S., Pedersen, G., Engen, M., Johansen, M. S., Johansson, P. N., Schlüter, C., ... and Bateman, A. W. (2013). 'The MBT Adherence and Competence Scale (MBT-ACS): Development, Structure and Reliability'. *Psychotherapy Research*, 23(6): 705–717. See also Muñoz Specht, P., Ensink, K., Normandin, L., and Midgley, N. (2016). 'Mentalizing Techniques Used by Psychodynamic Therapists Working with Children and Early Adolescents'. *Bulletin of the Menninger Clinic*, 80(4): 281–315.

a patient build trust in others in order to facilitate mentalization; a fourth may focus on helping a patient reduce non-mentalizing in order to reduce epistemic mistrust. All of these can legitimately claim to be mentalization-based approaches to therapy, given that the constructs of mentalization and epistemic trust both serve as switch-points for different and non-overlapping groups of meanings.

The latest version of the manual for delivering MBT was published in 2016.³⁹ It draws on the mature model of mentalizing and non-mentalizing (see Chapters 4 and 5), as well as elements of the account of the self (see Chapter 6). The latest manual also shifts from a deficit discourse about non-mentalizing towards a greater emphasis on the potentially adaptive basis and role of these modes of mental processing in the context of adversities (see Chapter 7). MBT now begins with a 10-week group-based programme, which comprises psychoeducation about mentalization and support for patients in establishing their expectations for therapy. This preliminary psychoeducation programme appears to be experienced by many patients as relevant and useful.⁴⁰ However some individuals do decide at this point to drop out, rather than initiating individual MBT or individual + group MBT.⁴¹ A recent study by Jørgensen reported that patients with BPD who reported low reflective functioning were at increased risk of dropping out of group MBT but not treatment as usual.⁴² Such findings agree with Bateman and Fonagy's suspicion that MBT, and especially group MBT, may require some threshold degree of mentalizing skills, and highlights the importance of the introductory psychoeducation to the extent that it can lay the ground for these skills.

Bateman and Fonagy proposed that individual therapy should begin with an assessment of the specific forms of mentalizing that appear to be areas of existing strengths and weaknesses, as part of the development of a shared formulation:

The clinician should draw up a mentalising profile. This involves locating the individual's style of mentalising on each of the different dimensions, and then considering the relationship between these different ways of functioning on each dimension: that is, do they cause mentalising difficulties to snowball, or do they compensate for each other?⁴³

³⁹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press.

⁴⁰ Ditlefsen, I. T., Nissen-Lie, H. A., Andenæs, A., Normann-Eide, E., Johansen, M. S., and Kvarstein, E. H. (2020). "Yes, there is actually hope!"—A qualitative investigation of how patients experience mentalization-based psychoeducation tailored for borderline personality disorder'. *Journal of Psychotherapy Integration*, Early View.

⁴¹ Barnicot and Crawford report a 64% drop-out rate by the end of the group-based programme. It is not clear how representative this is of other implementations of MBT. Barnicot, K. and Crawford, M. (2019). 'Dialectical Behaviour Therapy v. Mentalisation-Based Therapy for Borderline Personality Disorder'. *Psychological Medicine*, 49(12): 2060–2068. Other studies have tended to report attrition once individual MBT begins. When attrition is reported at all, rates vary between 7% (Brune, Dimaggio, and Edel, 2013) and 72% (Robinson et al., 2016): Brune, M., Dimaggio, G., and Edel, M. A. (2013). 'Mentalization-Based Group Therapy for Inpatients with Borderline Personality Disorder: Preliminary Findings'. *Clinical Neuropsychiatry*, 10: 196–201; Robinson, P., Hellier, J., Barrett, B., Barzdaitiene, D., Bateman, A., Bogaardt, A., . . . and Kern, N. (2016). 'The NOURISHED Randomised Controlled Trial Comparing Mentalisation-Based Treatment for Eating Disorders (MBT-ED) with Specialist Supportive Clinical Management (SSCM-ED) for Patients with Eating Disorders and Symptoms of Borderline Personality Disorder'. *Trials*, 17(1): 549.

⁴² Jørgensen, M. S., Bo, S., Vestergaard, M., Storebø, O. J., Sharp, C., and Simonsen, E. (2021). 'Predictors of dropout among adolescents with borderline personality disorder attending mentalization-based group treatment'. *Psychotherapy Research*, Early View. On dropout from MBT, see also Andersen, C. F., Poulsen, S., Fog-Petersen, C., Jørgensen, M. S., and Simonsen, E. (2020). 'Dropout from mentalization-based group treatment for adolescents with borderline personality features: A qualitative study'. *Psychotherapy Research*, Early View.

⁴³ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 115.

The weaknesses that should form the particular target of intervention should be those that contribute to such snowballing, as well as those that are especially pertinent to the patient, contributing to jointly agreed goals for the work together. Existing mentalizing strengths can be drawn upon as assets and resources in the therapy, though it should also be considered how they may be used as substitutes or compensation for aspects of mentalizing that have been left underdeveloped. Bateman and Fonagy indicated that mentalization should not be taken to be a unitary capacity, and that discrepancies between the different capabilities under the umbrella of mentalization are very much to be expected. A formulation of the patient's strengths and weaknesses in areas of mentalizing, and how these relate to current relationships and behaviour, should be written down and shared with the patient. It should then be subsequently revised periodically.

If the patient comes to therapy already with a diagnosis assigned to them, whether by a clinician or self-ascribed, the therapist should help the patient to specify the particular symptoms that cause them trouble, and the history of these symptoms. This will allow the therapist and patient to generate a developmental narrative about how the symptoms emerged. The diagnosis should not in itself be treated as an account or explanation of mental states.⁴⁴ Building on their recent thinking about adaptation (see Chapter 7), Bateman and Fonagy argued that strategic aspects of symptoms should be recognized before seeking to effect change, in order to minimize potential iatrogenic effects of intervention. In the psychoeducation that begins MBT, patients are taught about the conditional attachment strategies, and supported to reflect on the ways in which their behaviour when hurt, anxious, or distressed resembles one of these patterns. Patients are told that at times they may use one or the other conditional strategy, and sometimes they may feel that they use both together in relatively more or less coherent ways. Patients are not taught about the disorganized classification, though no reason is provided by Bateman and Fonagy for this absence from the curriculum. Perhaps they anticipate that the disorganized attachment classification would be experienced as disempowering by patients.

Bateman and Fonagy claimed that the key task for the MBT therapist is to help the patient balance the four poles of mentalizing: internal and external; affective and cognitive; self and other; implicit and explicit. But they make little of the need to balance explicit with implicit mentalizing, and in fact as we have seen 'active, intentional effort to find out about opaque mental life' is one of the defining principles of MBT. A balance between internal and external, and between self and other, translates primarily into two implications. A first is that the clinician should seek to ensure that patients are capable of both. A second is that patients should be capable of withdrawing from one to pursue the other as needed. Rather than 'balance', Bateman and Fonagy seem more concerned with the flexible use of capacities.

By contrast, the issue of balance is more clearly evident in relation to affect and cognition. This is because the overarching clinical model of Bateman and Fonagy is that reductions in symptoms arise from supporting patients during clinical sessions to simultaneously mentalize and retain emotional equilibrium. Whereas humanistic approaches to therapy generally regard expressions of empathy for the patient as beneficial, an MBT perspective treats such expressions as appropriate only when they will facilitate mentalization and not

⁴⁴ Addressing mental states merely as effects of diagnoses has been characterized as poor reflective functioning. Fonagy, P., Target, M., Steele, H. and Steele, M. (1998). *Reflective Functioning Manual*, Version 5, London: UCL/Anna Freud Centre: 'The use of diagnostic terminology, or reference to mental illness, should be considered very carefully, and on the whole rated low, if this is the sole explanation offered for the caregiver's behaviour, and the specific mental states of caregiver and other persons affected are not specified' (p. 21).

contribute to excess intensity or drama in the therapeutic relationship. In fact, when a patient is upset and this prompts forms of non-mentalizing, the MBT clinician should prioritize helping the patient regain their calm, so that interest can then be sparked in mental states. This is termed a ‘contrary move’. Such a move may entail somewhat matter-of-fact responses, a more cognitive angle on the problem, a focus outward to the minds of other people, or moving the patient away from his or her current focus:

The clinician moves emotionally closer to the patient during a session only to the point at which he/she judges the patient is on the verge of losing mentalizing. At this point, he/she moves back, distancing him/herself from the patient, to reduce the level of emotional arousal. Here we encounter a clinically significant paradox—just when the clinician would naturally move emotionally closer to the patient, we ask that he/she moves away.⁴⁵

This stance shapes the MBT approach to trauma. Many psychotherapeutic modalities regard discussion of past traumas as essential for increasing insight and present-day functioning. Several of Fonagy’s collaborators, such as Jon Allen, have made claims that move in this direction.⁴⁶ However, for their part, Bateman and Fonagy have stated that traumatic past experiences are given attention in MBT either only when this will support the patient to gain skills in retaining the capacity to mentalize when distressed, or when it will support the development of a self-representation facilitative of mentalization.⁴⁷ Compared with psychodynamic psychotherapy, in MBT ‘there should be . . . an avoidance of extensive discussion of past trauma, except in the context of reflecting on current perceptions of the mental states of maltreating figures and on changes in mental state from one’s past as a victim to one’s experience now.’⁴⁸

Bateman and Fonagy expressed concern that if a story about specific traumatic past experience is created to make sense of present-day symptoms—for instance, a demanding parent as the source of a patient’s contemporary anxiety—this can flatten the complexity of both the past and the present.⁴⁹ The result may stimulate pretend mode for patients, focusing attention on an origin myth rather than helping the patient to use attention to mental states

⁴⁵ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 195. On contrary moves, see also Bateman, A. (2019). ‘Mentalizing, Mentalization-Based Treatment and BPD’. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast/>.

⁴⁶ Allen, J. G. (2013). *Mentalizing in the Development and Treatment of Attachment Trauma*, London: Routledge.

⁴⁷ Fonagy, P. (2010). ‘Attachment Trauma and Psychoanalysis: Where Psychoanalysis Meets Neuroscience’, in M. Leuzinger-Bohleber, J. Canestri, and M. Target (eds). *Early Development and its Disturbances: Clinical, Conceptual and Empirical Research on ADHD and other Psychopathologies and its Epistemological Reflections*, London: Karnac Books, pp. 53–75: ‘I believe that reconstruction is essential to the therapeutic process because: i. it provides a means to bring the patient’s mind into contact with what it has previously found intolerable; ii. It provides a place where threat to the ego and therapeutic goal are reasonably balanced; iii. It generates a coherent self-narrative assuming a historical continuity of self which may itself be of therapeutic value (Holmes 1998; Shafer 1980; Spence 1994), and iv. Most importantly, it can help in the primary task of the recovery of mentalisation’ (p. 68).

⁴⁸ Fonagy, P., Bateman, A. W., and Luyten, P. (2012). ‘Introduction and Overview’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 3–42, p. 39.

⁴⁹ This clinical recommendation was already put forward by Bion, for whom too central a focus on a past traumatic event as the perceived origin of present difficulties ran the risk of ‘saturating’ the situation, blocking opportunities for further learning by giving the sense of explanatory satisfaction. Bion, W. R. (1970). *Attention and Interpretation*, London: Karnac Books, p. 29. See also Bion, W. R. ([1973] 1990). *Brazilian Lectures*, London: Karnac Books: ‘We bring into the open certain elements of an analysand’s past, not because we think they are particularly valuable, but because they are not valuable for him to have in his luggage. If we bring them to the surface then he can forget them. Those memories, past or future, which he does not know seem to have a great deal of power; they are what I would call feeble ideas but powerful emotions’ (p. 30).

to account for observable social behaviour and/or perceptual experience.⁵⁰ This is not to say that memories of trauma recalled during therapy should be regarded solely as constructions or fantasy. As we saw in Chapter 2, child patients whose therapists treated their accounts of abuse and neglect as fantasy had especially poor long-term outcomes in the Anna Freud Centre Retrospective study.⁵¹ Fonagy has reflected that traumas recalled in therapy can also serve as important ‘crystallization of an individual’s struggles with his or her circumstances, identity, relationships, and life’, and so may have importance for formulation.⁵² When a patient experiences something as a felt truth in a therapeutic context—and the impact of a past trauma can certainly serve this function—the effect can be to move the patient closer to the therapist and increase epistemic trust.⁵³

Nonetheless, too exclusive a focus on identifying or remembering a particular trauma as the source of current symptoms can work against the toleration of uncertainty necessary for supporting reflection on mental states and the different perspectives they entail.⁵⁴ More generally, Bateman and Fonagy have argued, to inhibit entry into hypermentalizing or other forms of pretend mode ‘the clinician should avoid inadvertently creating situations where the patient is forced into talking about mental states that they cannot immediately link to subjectively felt reality.’⁵⁵ Instead, the therapist should aim to keep conversations centred around whatever is felt to be current and emotionally salient ‘in the patient’s mind, in other words, in working memory.’⁵⁶ As we will see in Chapter 9, Fonagy and colleagues have very recently questioned whether an individual’s working memory is the only site of mental states, or whether groups or institutions can also have beliefs and atmospheres that can be the target of mentalizing. However, in Bateman and Fonagy’s 2016 manual for MBT, it is the individual alone who can be the subject of mental states and the target of mentalizing.

One of the essential techniques advocated by Bateman and Fonagy is what they have called ‘mentalizing functional analysis’. In light of the clinician’s understanding of the patient’s strengths and weaknesses, the MBT therapist keeps a look out for times when the patient seems to drop out of mentalizing and enter into a non-mentalizing mode of mental processing. At that point, the clinician can ask the patient to ‘stop and re-wind’, and narrate the

⁵⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: ‘MBT has never been focused on the past or been insight oriented, although it is respectful of past experience and how it affects the present. It is our view that explaining a person’s current wish to please in terms of a continuing wish to satisfy a demanding image of a parent is a descriptor masquerading as an explanation. Furthermore, it potentially has the harmful side effect of stimulating pretend mode’ (p. 218).

⁵¹ See also Daly, E. (1997). ‘Women from a Broken Home’, Independent, 21 May. Accessed at: <http://www.independent.co.uk/life-style/women-from-a-broken-home-1262770.html>: ‘Bernice Andrews, an undoubted opponent of the Society, believes there is strong evidence—not least from a survey she conducted—that patients can recover memories. She also believes some such memories are false. So does Professor Peter Fonagy, who has resigned from the advisory board of the British False Memory Society because, he says, “the more recent evidence that has been coming through . . . has been somewhat inconsistent with the position of the False Memory Society.”’

⁵² Fonagy, P. (2016). ‘Ask the Experts: A Conversation with Peter Fonagy’, *Trauma Matters*, Spring, p. 3. Accessed at: https://docs.wixstatic.com/ugd/8d6d78_d51caa6cb0554a68bb39e0bca78eaa57.pdf: As a clinician, ‘you will inevitably encounter traumatic experiences. What you find there is not necessarily the cause of an individual’s problems, but it is inevitably a crystallization of an individual’s struggles with his or her circumstances, identity, relationships, and life’ (p. 3).

⁵³ Allison, E. and Fonagy, P. (2016). ‘When is truth relevant?’ *Psychoanalytic Quarterly*, 85(2): 275–303, p. 294.

⁵⁴ Fonagy, P. (2003). ‘Rejoinder to Harold Blum’, *The International Journal of Psychoanalysis*, 84(3): 503–509: ‘While reconstruction of how things actually were in childhood can significantly contribute to therapeutic action, it is the process rather than the outcome of this reconstruction that is therapeutic, due to the opportunity thus afforded to rework current experiences in the context of other perspectives.’ (p. 503).

⁵⁵ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 208.

⁵⁶ Ibid. 218–219.

sequence again, so that the clinician and patient can work collaboratively to identify when and how mentalizing was lost and non-mentalizing began.⁵⁷ Special attention is paid to the feelings that may stimulate forms of non-mentalizing. The therapist and patient may then reflect together about what strategies could be implemented in future to halt the decay of mentalizing or the advance of non-mentalizing. Bateman and Fonagy recognized, however, that clinicians too may enter non-mentalizing states. For instance, they may find themselves just going along with what is being said, having entered into pretend mode. They may find themselves caught up in feelings that seem so pressing that new information cannot enter, having entered a state of psychic equivalence. They may find their minds dominated by an urgent wish to do something constructive, having entered teleological mode.⁵⁸ When the patient or clinician perceives that the clinician has started to non-mentalize, Bateman and Fonagy likewise advised use of ‘stop and re-wind’.

‘Stop and re-wind’ was developed as a clinical technique, but it also has theoretical stakes. Over the decades, Fonagy and collaborators have given less and less explicit notice to attentional processes in their thinking about mentalization; it seems to have dropped, unceremoniously, out of sight.⁵⁹ Attentional processes are not described by Bateman and Fonagy as part of what MBT attempts to change. Nonetheless, ‘stop and re-wind’ can be considered above all as a technique for practising both attentional flexibility and attentional focus as component parts and contributors to mentalizing.

This is all the more surprising because Fonagy and Gergely both attribute the basis of the idea of ‘ostensive cues’ to Sperber and Wilson, for whom such cues are primarily a kind of way of directing the attention of a conversation partner. Sperber appears even to somewhat regret that in highlighting ostension he appears to have drawn away recognition from the direction of others’ attention as the underlying and wider issue.⁶⁰ In the work of Fonagy and colleagues, concern with ostensive cues has rather supplanted the wider question of attentional processes, which they had considered in the 1990s and early 2000s. In the 2016 MBT manual, Bateman and Fonagy and colleagues highlighted the importance of ostensive cues provided by the therapist to the patient, and of validating the salience and relevance of the patient’s perspective as critical to the initiation of any further discussion of that perspective. For instance, marked mirroring is a powerful ostensive cue: conveying, in the interplay of

⁵⁷ Ibid. 228.

⁵⁸ Ibid. 284.

⁵⁹ A study by Perroud and colleagues, including Fonagy, examined self-reported mentalization and attention deficit hyperactivity disorder (ADHD). However, this paper addressed attention difficulties solely as an effect of mentalization, not as a contributing factor or component of mentalizing, as in Fonagy’s earlier work. Perroud, N., Badoud, D., Weibel, S., Nicastro, R., Hasler, R., Küng, A. L., ... and Prada, P. (2017). ‘Mentalization in Adults with Attention Deficit Hyperactivity Disorder: Comparison with Controls and Patients with Borderline Personality Disorder’. *Psychiatry Research*, 256: 334–341. Curiously, the Swiss authors, in a paper without Fonagy as co-author, did later consider attention as a contributor to mentalizing, developing a model of bidirectional influence. Badoud, D., Rüfenacht, E., Debbané, M., and Perroud, N. (2018). ‘Mentalization-Based Treatment for Adults with Attention-Deficit/Hyperactivity Disorder: A Pilot Study’. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 21(3): 149–154. A concern with attention has been re-integrated with mentalization in the AMAR Model proposed by Lecannelier, F. (2019). ‘A Transcultural Model of Attachment and Its Vicissitudes: Interventions Based on Mentalization in Chile’. In *Clinical Handbook of Transcultural Infant Mental Health*. New York: Springer, pp. 135–149.

⁶⁰ E.g. Sperber, D. (2019). ‘Personal Notes on a Shared Trajectory’, in Kate Scott, Billy Clark, and Robyn Carston (eds), *Relevance, Pragmatics and Interpretation*, Cambridge: Cambridge University Press, pp. 13–20. See also recent blog posts by Sperber, accessed at <http://cognitionandculture.net>. A conceptualization of the direction of attention as fundamental to ostension and to guiding meaning-making appears already in Sperber’s earliest works. For instance, Sperber, D. (1975). *Cultural Symbolism*, Cambridge: Cambridge University Press: ‘cultural symbolism focusses the attention of the members of a single society in the same directions, determines parallel evocational fields that are structured in the same way, but leaves the individual free to effect an evocation in them as he likes. Cultural symbolism creates a community of interest but not of opinion’ (p. 137).

primary and secondary meanings, both that the therapist acknowledges the patient's emotion and intentionality, and that the therapist can cope with these responses. Use of these techniques is assumed to 'increase the patient's epistemic trust and thus acts as a catalyst for therapeutic success'.⁶¹ While these techniques are common to many therapeutic modalities, Fonagy and colleagues anticipated that MBT could give them particular weight. In fact, they might even form the basis in the future for forms of MBT delivered over the internet, because Fonagy and colleagues regard ostensive cues rather than therapeutic alliance as reigniting epistemic trust and the capacity to learn from experience.⁶²

This concern with ostensive cues has also led to theoretical revision. As we saw in the previous chapter, Fonagy and colleagues had initially described ostensive cues as behaviours that acknowledge the i) mental states; ii) intentionality; and iii) individuality of the subject. Initially, these properties of ostensive cues were extrapolated from observations of infants to later development—for instance, to adult practices that entail mutual recognition of intent. However, in a 2019 chapter, Fonagy, Allison, and Campbell have argued that, after infancy, the perception of 'understanding of the individual's personal narrative by another person creates a potential for epistemic trust'.⁶³ Above all, 'recognition of subdominant narratives is a particularly potent way of establishing epistemic trust'.⁶⁴ They define personal narrative as a person's 'imagined sense of self' at a given moment, and identify that the clarity and specificity of this narrative is in significant part a function of mentalizing of the self.⁶⁵ The imagined sense of self is nonetheless complex and heterogenous, inflected by a person's developmental history, their cultural context, the pragmatic demands of their immediate circumstances, alien intentions from the psychoanalytic unconscious, and sometimes significant

⁶¹ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9.

⁶² Fonagy, P. and Clark, D. M. (2015). 'Update on the Improving Access to Psychological Therapies Programme in England: Commentary on ... Children and Young People's Improving Access to Psychological Therapies'. *BJPsych Bulletin*, 39(5): 248–251: 'The remarkable success of internet-delivered therapies strongly challenges the claim that a strong therapeutic alliance is essential' (p. 249). See also Falconer, C. J., Cutting, P., Davies, E. B., Hollis, C., Stallard, P., and Moran, P. (2017). 'Adjunctive Avatar Therapy for Mentalization-Based Treatment of Borderline Personality Disorder: A Mixed-Methods Feasibility Study'. *Evidence-Based Mental Health*, 20(4): 123–127. At best, they regard the concept of 'therapeutic alliance' as an umbrella term that points weakly at the kinds of interaction facilitative of epistemic trust. Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast/>.

⁶³ Fonagy, P., Allison, E., and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, p. 70.

⁶⁴ Fonagy, P., Campbell, C. and Allison, E. (2019). 'Therapeutic models', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 171. The turn to attention to narrative may reflect Liz Allison's particular interests in narrative and storytelling. E.g. Allison, E. (2017). 'Observing the Observer: Freud and the Limits of Empiricism'. *British Journal of Psychotherapy*, 33(1): 93–104.

⁶⁵ Fonagy, P., Allison, E. and Campbell, C. (2019). 'Mentalising, Resilience and Epistemic Trust', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 63–77, pp. 72 and 74; Fonagy, P. and Allison, E. (2018). 'The Origin of Human Life: A Psychoanalytic Developmental Perspective', European Psychoanalytical Federation, 31st Annual Conference, 24 March, Warsaw, 'It is perhaps helpful here to bear in mind a distinction that Sandler made between two different levels at which the concept of self-representation could be considered. On the one hand, he thought that the self-representation exists as a more or less enduring organization, schema or set of rules that is constructed out of a multitude of impressions and exists outside subjective experience even though it is in large part a consequence of subjective experience. On the other hand, there is the level of the phenomenal or experiential self-representation, which is "the subjective representation or subjective experience of ourselves that we have, at any given moment, in what we can refer to as the experiential realm." It is this second sense that we wish to capture with the term personal narrative.'

involvement of the primary unconscious (see Chapter 6). There is considerable scope, therefore, for subdominant threads that are knowable by an individual and by others, but that do not ordinarily feature as part of an individual's espoused self-representation.⁶⁶ Epistemic trust is facilitated by the felt acuity of the match between some part of our self-representation and the account of us offered by another. A lack of clarity and specificity in this narrative will render individuals vulnerable to missing accurate matches (epistemic hypervigilance), or incorrectly identifying accurate matches (epistemic credulity), or both (see Chapter 7).

The characterization of others' perceived understanding one's personal narrative as an ostensive cue is innovative. It is not implied by the previous concept of ostensive cues as behaviours that acknowledge the i) mental states; ii) intentionality; and iii) individuality of a subject. The 'imagined sense of self' depicted by a personal narrative is not the sum of thoughts and feelings, though it plays a role in generating them and partly takes its content from them. Nor is it merely the intentionality of the subject, though this may play a part. The claim that epistemic trust is generated by recognition of personal narrative rather than acknowledgement of mental states and intentionality appear to signal a wish by Fonagy and colleagues, not as yet fully pursued, to bring to bear more of the complex account of the self on matters of clinical technique. Given Fonagy's criticisms of the Bakhtin-inspired model of dominant and subdominant personal narratives in CAT in the 1990s (see Chapter 6), it would be interesting to see how Fonagy and colleagues would now regard this model of personal narratives and their role in clinical technique.⁶⁷ This would be of particular note given Fonagy and colleagues' growing interests in culture. The Bakhtinian approach in CAT holds that personal narrative is woven out of cultural resources. From a mentalizing perspective, such resources may suggest or impose codes that relatively hinder or facilitate the construction, identification, modulation, and expression of mental states.

Fonagy and colleagues have been continually active in evolving and developing MBT. Meanwhile, trials have been accumulating that have assessed the effectiveness of the modality in its different incarnations. Most of these, though certainly not exclusively, have focused on its application with adult patients with BPD. True, the relationship between mentalizing difficulties and BPD may be less tight than initially anticipated by Fonagy. For instance, Bouchard, Target, Lecours, Fonagy, and colleagues have found only an association of $r = -.24$ between symptoms of personality disorder and reflective function scored on the Adult Attachment Interview.⁶⁸ Nonetheless, the importance of mentalizing difficulties has

⁶⁶ Horne, among others, has previously criticized the tendency of Fonagy and colleagues to give insufficient attention to subdominant narratives. Horne, M. (2004). 'The Universe of Our Concerns: The Human as Person in the Praxis of Analysis.' *Journal of Analytical Psychology*, 49(1): 33–48. There are some points of similarity between these recent remarks by Fonagy, Allison, and Campbell on acknowledgement of subdominant narratives as an ostensive cue, and earlier work by Feldman, M. (2007). 'Addressing Parts of the Self'. *The International Journal of Psychoanalysis*, 88(2): 371–386. Though, unlike Fonagy and colleagues, Feldman also considers how recognition of subdominant narratives may precisely disturb epistemic trust, through being experienced as an intrusion or threat to the dominant self-representation.

⁶⁷ Pollard, R. (2008). *Dialogue and Desire: Mikhail Bakhtin and the Linguistic Turn in Psychotherapy*, London: Karnac Books. It may also provide impetus for Fonagy and colleagues to clarify whether they regard 'thinking' as necessarily verbal-discursive, and more generally help situate the boundaries of what is meant by 'thought' as a 'mental state' compared with the building blocks of thoughts.

⁶⁸ Bouchard, M. A., Target, M., Lecours, S., Fonagy, P., Tremblay, L. M., Schachter, A., and Stein, H. (2008). 'Mentalization in Adult Attachment Narratives: Reflective Functioning, Mental States, and Affect Elaboration Compared'. *Psychoanalytic Psychology*, 25(1): 47–66. Another study found no association between reflective function and the severity of personality disorder symptoms. Fischer-kern, M., Schuster, P., Kapusta, N. D., Tmej, A., Buchheim, A., Rentrop, M., ... and Fonagy, P. (2010). 'The Relationship between Personality Organization, Reflective Functioning, and Psychiatric Classification in Borderline Personality Disorder'. *Psychoanalytic Psychology*, 27(4): 395–409. Researchers have also found a very weak ($r = -.11$) association between Theory of Mind and BPD symptoms. Belsky, D. W., Caspi, A., Arseneault, L., Bleidorn, W., Fonagy, P., Goodman, M., ... and Moffitt, T. E. (2012). 'Etiological Features of Borderline Personality Related Characteristics in a Birth Cohort

been indicated by the finding that, in the large majority of studies, patients with personality disorders benefit from MBT and show fewer symptoms and better day-to-day functioning by the end of the trial. The same is also generally true of other comparison interventions when patients have BPD alone, without other conditions. Fonagy and colleagues have interpreted this finding with a speculative proposal: that even if the models are different, the use of any evidence-based model by therapists gives coherence to the way the patient is recognized as an intentional agent by the therapist, which serves as an ostensive cue and facilitates epistemic trust.⁶⁹

However, for patients with complex personality disorders, MBT has proven the most effective psychological modality.⁷⁰ This may imply that the difficulties with mentalizing identified initially by Fonagy and colleagues as especially characteristic of BPD are in fact less specific. This would seem supported by the Sharp and colleagues' study from 2015 mentioned in the previous chapter, in which BDP loaded with the general p-factor. To untangle these questions, it would be necessary to understand the relative role of mentalization in mediating clinical and functional outcomes for patients, and whether this differs between treatment modality. Fonagy and colleagues have argued that mentalization is improved by many talking therapies, and should be treated as the central target for effecting therapeutic change. In support for this claim, De Meulemeester found that improvement on the Reflective Functioning Questionnaire scale for indiscriminate uncertainty (RFQu) among patients receiving hospital-based psychoanalytic treatment was associated to the highest degree with relief from symptoms ($r = .89$).⁷¹ The remarkable strength of this association is such that, if replicated, it suggests that RFQu and relief from symptoms may even be part

of 12-Year-Old Children'. *Development and Psychopathology*, 24(1): 251–265, Figure 1a. On the interpretation of effect sizes in psychology, see Funder, D. C. and Ozer, D. J. (2019). 'Evaluating Effect Size in Psychological Research: Sense and Nonsense'. *Advances in Methods and Practices in Psychological Science*, 2(2): 156–168.

⁶⁹ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9. In fact, this claim warrants significant further scrutiny. Is it the *accuracy* of the model that is contributing to the patient's sense of feeling recognized? Or the clinician's *confidence* in the model? Or the *coherence of the narrative* the model permits the therapist and/or permits the patient? It does not seem fully plausible that Kleinian psychoanalysis and exposure therapy, though both could claim to have supporting evidence, have the same approach to recognizing the patient as an intentional agent. Or if recognition is achieved by both through some generic process (e.g. therapist belief that the patient has intentions and mental states; use of ostensive cues; adequate professional boundaries), why this recognition would not be achieved by non-evidence-based therapies.

⁷⁰ Fonagy, P., Luyten, P., and Bateman, A. (2017). 'Treating Borderline Personality Disorder with Psychotherapy: Where Do We Go From Here?'. *JAMA Psychiatry*, 74(4): 316–317. 'The various evidence-based treatments included in the present meta-analysis may have studied different populations and thus may be differentially effective in different subtypes of BPD. For example, mentalization-based therapy was superior over control (structured clinical management) only in patients with BPD with multiple Axis II diagnoses' (p. 317). See also Volkert, J., Hauschild, S., and Taubner, S. (2019). 'Mentalization-Based Treatment for Personality Disorders: Efficacy, Effectiveness, and New Developments'. *Current Psychiatry Reports*, 21(4): 25; Kvarstein, E. H., Pedersen, G., Folmo, E., Urnes, Ø., Johansen, M. S., Hummelen, B., ... and Karterud, S. (2019). 'Mentalization-Based Treatment or Psychodynamic Treatment Programmes for Patients with Borderline Personality Disorder—The Impact of Clinical Severity'. *Psychology and Psychotherapy: Theory, Research and Practice*, 92(1): 91–111.

⁷¹ De Meulemeester, C., Vansteelandt, K., Luyten, P., and Lowyck, B. (2018). 'Mentalizing as a Mechanism of Change in the Treatment of Patients with Borderline Personality Disorder: A Parallel Process Growth Modeling Approach'. *Personality Disorders: Theory, Research, and Treatment*, 9(1): 22. Though not an intervention study, Morosan documented in a community-based cohort study of adolescents that self-reported indiscriminate uncertainty about mental states (RFQu) was associated with more externalizing behaviours at the start of the study; reductions in self-reported certainty about mental states (RFQc) was associated with sharper declines in externalizing behaviours over time. Morosan, L., Ghisletta, P., Badoud, D., Toffel, E., Eliez, S., and Debbané, M. (2019). 'Longitudinal Relationships between Reflective Functioning, Empathy, and Externalizing Behaviors During Adolescence and Young Adulthood'. *Child Psychiatry & Human Development*, 51: 59–70.

of the same underlying construct. However, to the surprise of the researchers, there was no link with the RFQ certainty scale, though, in general, the RFQc scale has not performed as anticipated across several studies (see Chapter 7). It is also notable that a major recent meta-analysis found that mentalization-based parenting interventions result in little change in reflective functioning.⁷² Such findings raise the question of the extent to which mentalization in general is indeed the mechanism of change as anticipated by Fonagy, or whether further refinement might be needed in specifying and measuring how MBT is anticipated to work, perhaps with greater focus on indiscriminate uncertainty about mental states.

As evidence that the effectiveness of MBT is mediated by mentalization, Fonagy and colleagues have frequently referred to a trial of MBT for self-harming adolescents by Rossouw.⁷³ To measure mentalization, Rossouw and Fonagy used the 'How I Feel' Questionnaire, which is a self-report assessment of emotion regulation. Participants report on the frequency, intensity, and control of feelings such as sadness, fear, and anger. It would be expectable that there would be shared variance between the 'How I Feel' Questionnaire with self-reported self-harm. For self-reported emotion regulation to partially mediate associations between treatment and self-reported self-harm cannot be regarded as proof that mentalization mediated the effectiveness of the intervention. Furthermore, the model tested included self-reported attachment avoidance as an additional mediator, making the mediation by self-reported emotion regulation difficult to interpret with confidence. The researchers did not report whether scores on How I Feel mediated effects of the intervention on symptoms of BPD or on depression, though both were measured, and such tests were indicated by Rossouw's own developmental model of mediating factors.⁷⁴ The Rossouw study also has among the highest attrition rates of any trial of MBT published to date, and no assessment was made of whether participants who left the study differed from participants who remained in therapy to completion. Overall, the Rossouw study cannot be treated as evidence of the role of mentalizing in mediating the effects of MBT on outcomes. This should be regarded as still yet to be tested.

The potential for harm

One outstanding question faced by MBT is the potential for harm to patients stemming from the therapy—iatrogenesis. Fonagy and colleagues have been vocal about the potential for other treatment modalities to cause harm to patients. As we saw in Chapter 1, Fonagy had particular concern with the ways that classical psychoanalytic technique may at times actively contribute to problems with emotion modulation for patients with mentalizing difficulties.⁷⁵ Fonagy and colleagues have in principle acknowledged the potential for iatrogenesis in mentalization-based therapies. But this has generally been only with passing references, and they have not developed these reflections into hypotheses for testing.⁷⁶ The

⁷² Lo, C. K. and Wong, S. Y. (2020). 'The effectiveness of parenting programs in regard to improving parental reflective functioning: A meta-analysis'. *Attachment & Human Development*, Early View.

⁷³ Rossouw, T. I. and Fonagy, P. (2012). 'Mentalization-Based Treatment for Self-Harm in Adolescents: A Randomized Controlled Trial'. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(12): 1304–1313.

⁷⁴ Rossouw, T. I. (2012). *Self-Harm in Young People: Randomised Control Trial Testing Mentalisation Based Treatment against Treatment As Usual*. Unpublished MD thesis, London: University College London. Diagram 2, p. 34.

⁷⁵ E.g. Fonagy, P. and Bateman, A. (2006). 'Progress in the Treatment of Borderline Personality Disorder'. *British Journal of Psychiatry*, 188: 1–3, p. 1.

⁷⁶ In papers from 1992 and 1994, Fonagy and colleagues acknowledged that improved capacities for mentalizing will allow an individual to 'conceive of his world in new, sometimes sadder and sometimes happier ways'. Fonagy,

sole apparent exception to the lack of explicit hypothesis generation about the potential for iatrogenesis from MBT is a proposal from 2005 by Fonagy and Target. They stated that declines in use of psychic equivalence lead to an 'almost inevitable worsening of the patient's symptoms' so long as mentalizing skills remain weak, because the experience of personal coherence permitted by psychic equivalence is initially undermined.⁷⁷ However, this hypothesis has not been operationalized and tested, or even revisited. In part, this is likely to have been due to the available measures, which generally treat non-mentalizing as the opposite of mentalizing, and so are not able to characterize a state in which i) non-mentalizing has declined but ii) mentalizing has yet to increase.⁷⁸

Moving from consideration of theory to research findings, one empirical report suggesting a price attached to mentalizing appeared in the Anna Freud Centre retrospective study (Chapter 2). Using the Adult Attachment Interview, the researchers found that 'earned secure people may be reflective and functional, but they tend to be somewhat more vulnerable to depression than their unearned secure peers. The children who are most reflective about their families are also more worried about them.'⁷⁹ Other results suggestive of iatrogenesis have started to accumulate. For instance, Suchman and colleagues found that strengths in mentalizing the self in substance-using mothers were positively associated with depression ($r = .41$). Strengths in mentalizing their child ($r = .25$) were also associated with depression.⁸⁰ The study was cross-sectional so it is difficult to interpret causality, but it may be hypothesized that greater awareness of mental states is painful for substance-using parents. In another study, Stacks and colleagues likewise found a positive association between mentalizing

P., Moran, G. S., and Target, M. (1992). 'Aggression and the Psychological Self'. *Bulletin of the Anna Freud Centre*, 15: 269–284, p. 282. Repeated in Fonagy, P., Steele, M., Steele, H., Higgitt, A., and Target, M. (1994). 'The Emanuel Miller Memorial Lecture 1992: The Theory and Practice of Resilience'. *Journal of Child Psychology and Psychiatry*, 35(2): 231–257, p. 251. Likewise, the phrase 'mind reading may not be an unequivocally positive experience' first appeared in Fonagy, P., Steele, H., Steele, M. and Holder, J. (1997). 'Attachment and Theory of Mind: Overlapping Constructs?'. *Association for Child Psychology and Psychiatry Occasional Papers*, 14: 31–40, p. 37. It then appeared in several subsequent papers. However, in none of these cases was the passing phrase elaborated. In a recent paper, Barbara Castro Batic and Daniel Hayes at UCL/Anna Freud Centre have developed hypotheses for testing regarding potential mechanisms through which harm can occur to patients in talking therapies. However they do not attend to MBT specifically. Castro Batic, B., and Hayes, D. (2020). 'Exploring harm in psychotherapy: Perspectives of clinicians working with children and young people'. *Counselling and Psychotherapy Research*, 20(4): 647–656.

⁷⁷ Fonagy, P. and Target, M. (2005). 'Some Reflections on the Therapeutic Action of Psychoanalytic Therapy', in J. Auerbach, K. Levy, and C. E. Shaffer (eds), *Relatedness, Self-Definition and Mental Representation: Essays in Honor of Sidney J. Blatt*, New York: Taylor & Francis, pp. 191–212, p. 207.

⁷⁸ A new measure would be able to test this hypothesis: Gagliardini, G., and Colli, A. (2019). 'Assessing Mentalization: Development and Preliminary Validation of the Modes of Mentalization Scale'. *Psychoanalytic Psychology*, 36(3): 249–258.

⁷⁹ Target, M. (2008). 'Commentary', in F. N. Busch (ed.), *Mentalization: Theoretical Considerations, Research Findings, and Clinical Implications* (Psychoanalytic Inquiry Book Series, Volume 29), New York: Analytic Press, pp. 261–279, p. 277. See also Mark, I. L., IJzendoorn, M. H. V., and Bakermans-Kranenburg, M. J. (2002). 'Development of Empathy in Girls during the Second Year of Life: Associations with Parenting, Attachment, and Temperament'. *Social Development*, 11(4): 451–468. 'Radke-Yarrow et al. (Radke-Yarrow, Zahn-Waxler, Richardson, Susman, and Martinez, 1994) emphasized the importance of recognizing interacting influences of both parent and child. The results of their study with 24–48-month-old children showed how important these interactions may be. The highest frequencies of empathic responses were from children with severely depressed mothers, problems of affect regulation, and a secure attachment relationship with their mother, whereas children of well or less severely depressed mothers with secure attachment relationships and without problems of affect regulation scored neither extremely high nor extremely low. It might be that these middle scores were actually the more optimal scores, and that the very high scores of the children of severely depressed mothers reflected the caregiving behavior that children with disorganized attachments display as toddlers' (p. 453).

⁸⁰ Suchman, N. E., DeCoste, C., Leigh, D., and Borelli, J. (2010). 'Reflective Functioning in Mothers with Drug Use Disorders: Implications for Dyadic Interactions with Infants And Toddlers'. *Attachment & Human Development*, 12: 567–585.

scores and symptoms of depression among mothers with childhood maltreatment histories ($r = .29$).⁸¹ Such findings raise questions about whether MBT has the potential for iatrogenesis related to depression for patients with historic or concurrent vulnerabilities.

An obstacle to understanding the potential for iatrogenesis from MBT has been the multiple and diffuse ways in which the term ‘mentalizing’ has been used by Fonagy and colleagues. In light of the specifications in Chapter 4, passing remarks by Fonagy and colleagues may be drawn together to suggest that MBT has three potential bases for iatrogenesis. A first is the potential for mentalization to make individuals more open and vulnerable to harmful environments. This will be dealt with at length in the next chapter. A second potential basis for iatrogenesis has been the particular concern of Target, perhaps stemming from her reflections on the Anna Freud Centre Retrospective Study. For instance, Rizq and Target have expressed concern that ‘where high levels of RF [reflective functioning] tip over into anxious and depressive ruminations, they may unhelpfully sustain a preoccupation with the self, rather than with another’s experience and needs’. This ‘suggests that personal therapy for some may sponsor an unhelpful dwelling on inter and intra-personal dynamics.’⁸² Fonagy has agreed: ‘The idea of long-term deep introspection, of really studying your own mind, might be good for the training of therapists. Is it helpful for people with mental health problems? I think it probably does harm.’⁸³ The notion of mentalization as a ‘focus’ on mental states appears in some definitions, particularly towards the end of the 2000s.⁸⁴ However, as we saw in Chapter 3, focus on mental states—mind-mindedness—is perhaps best regarded as a potential correlate of mentalization, rather than a necessary element.

Fonagy and colleagues have acknowledged that mind-mindedness ‘is likely to be one of those parental attributes that is most adaptive in moderation.’⁸⁵ The potential for improved mentalizing may prompt anxious and depressive ruminations, whether about the self or about others. This may be in the form of hypermentalization, in which the content of the ruminations is speculative. However, the rumination could just as well be grounded in accurate attention to the specifics of present or past perceptual experience.⁸⁶ The problem lies

⁸¹ Stacks, A. M., Muzik, M., Wong, K., Beeghly, M., Huth-Bocks, A., Irwin, J. L., and Rosenblum, K. L. (2014). ‘Maternal Reflective Functioning among Mothers with Childhood Maltreatment Histories: Links to Sensitive Parenting and Infant Attachment Security’. *Attachment & Human Development*, 16: 515–533.

⁸² Rizq, R. and Target, M. (2010). ‘If that’s what I need, it could be what someone else needs’. Exploring the role of attachment and reflective function in counselling psychologists’ accounts of how they use personal therapy in clinical practice: a mixed methods study’. *British Journal of Guidance & Counselling*, 38(4): 459–481, p. 475.

⁸³ Fonagy, P. (2015). ‘Peter Fonagy on Psychoanalysis and IAPT’. The History of Emotions Blog, posted on 14 May by Jules Evans. Accessed at: <https://emotionsblog.history.qmul.ac.uk/2015/05/peter-fonagy-on-psychoanalysis-and-iapt>. See also Lucassen, N., Tharner, A., Prinzie, P., Verhulst, F. C., Jongerling, J., Bakermans-Kranenburg, M. J., ... and Tiemeier, H. (2018). ‘Paternal History of Depression or Anxiety Disorder and Infant–Father Attachment’. *Infant and Child Development*, 27(2): e2070: ‘Father’s history of depression or anxiety disorder was not significantly related to infant–father attachment security in the total sample. Interestingly, daughters of fathers with a history of depression or anxiety had higher scores on attachment security than daughters of fathers without this diagnosis’ (p. 1); ‘A potential explanation for the unexpected direction of this effect is that fathers with a history of depression or anxiety might have developed better mentalizing capacities and are therefore better able to “read” their child; potentially increased by experiences of psychotherapy’ (p. 7).

⁸⁴ E.g. Bateman, A. and Fonagy, P. (2008). ‘Comorbid Antisocial and Borderline Personality Disorders: Mentalization-Based Treatment’. *Journal of Clinical Psychology*, 64(2): 181–194: ‘Mentalizing simply implies a focus on mental states in oneself or in others, particularly in explanations of behaviour’ (p. 182).

⁸⁵ Fonagy, P., Bateman, A. W., and Luyten, P. (2012). ‘Introduction and Overview’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 3–42, p. 12.

⁸⁶ See e.g. Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). ‘Depression’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418: ‘Alongside the major evolutionary advantages, there is a potential shadow side to mentalisation. First, self-awareness and self-consciousness bring with them social emotions such as embarrassment, shame, and

not in the inaccuracy of the mind-mindedness, but in the lack of the qualities of tentativeness and generative uncertainty, and in the ability to use mentalizing as a capacity, i.e. to turn it to simmer when it is not helpful.⁸⁷ Within publications associated with MBT, Fonagy and colleagues offer guidance for responding to hypermentalizing and promoting mentalization. However, they do not attend to the potential for the mind-mindedness prompted by mentalization to contribute to realistic forms of rumination.

A third particular threat for iatrogenesis lies in the multiple uses of insight into motivations and intentions. In the early 2000s, Judy Dunn and colleagues had found that skills in social cognition were associated with heightened sensitivity to criticism among young children. Sutton and colleagues also documented strengths in social cognition among ring-leader bullies.⁸⁸ Citing these findings in a paper from 2006, Fonagy and Target observed that 'the possession of the capacity to mentalize is neither a guarantee that it will be used to serve pro-social ends, nor a guarantee of protection from malign interpersonal influence. The acquisition of the capacity to mentalize may, for example, open the door to more malicious teasing, increase the individual's sensitivity to relational aggression, or even mean that they take a lead in bullying others.'⁸⁹ Though Fonagy and Target framed this as a concern about acquisition of the capacity to mentalize in general, in fact the research they are citing is about social cognition and theory of mind. The problem would appear to lie, not in mentalization in general, but specifically in the multiple uses of skills in interpreting the thoughts and feelings of others.

Bateman and Fonagy reflected further on these matters in thinking about the delivery of MBT for patients with ASPD. They observed that ASPD can be associated with strengths in understanding the thoughts and motivations of others, despite difficulties in mentalizing the self, and particularly their own feelings.⁹⁰ Bateman and Fonagy offered the conjecture that 'the sacrifice of certain types of mentalizing frees up capacity for other domains to develop.'⁹¹ This potential for trade-offs between cultivations of kinds of mentalization is important, though it has been masked by reification of mentalization as a unitary process, and inattention to the kinds of motivation and fantasy that prompt forms of mentalizing and non-mentalizing. Bateman and Fonagy encouraged the rebalancing of mentalization in cases where there have been trade-offs, but do not consider the specific potential for iatrogenesis when rebalancing proves challenging. They also do not consider the role of

guilt ... awareness of being unable to achieve one's aspirations may lead to feelings of depression, loss, pain and fatigue' (p. 396).

⁸⁷ Asen, E. and Fonagy, P. (2017). 'Mentalizing Family Violence. Part 2: Techniques and Interventions'. *Family Process*, 56(1): 22–44: "Effective" mentalizing is not the uninterrupted capacity to be reflective and to mentalize explicitly at all times: This would not only be completely unsustainable, but would also kill spontaneity. Instead, the aim of therapy is to establish mentalizing in a balanced way which involves all family members and adapts flexibly and creatively to the context as and when needed' (p. 27).

⁸⁸ Cutting, A. L. and Dunn, J. (2002). 'The Cost of Understanding Other People: Social Cognition Predicts Young Children's Sensitivity to Criticism'. *Journal of Child Psychology & Psychiatry*, 43: 849–860. Sutton, J., Smith, P. K., and Swettenham, J. (1999). 'Social Cognition and Bullying: Social Inadequacy or Skilled Manipulation?'. *British Journal of Developmental Psychology*, 17(3): 435–450.

⁸⁹ Fonagy, P. and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology'. *Journal of Personality Disorders*, 20: 544–576, p. 559. The passage is then repeated in Fonagy, P., Gergely, G., and Target, M. (2007). 'The Parent–Infant Dyad and the Construction of the Subjective Self'. *Journal of Child Psychology and Psychiatry*, 48(3–4): 288–328, p. 307.

⁹⁰ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 56.

⁹¹ *Ibid.* 59.

uncivilized motivations and fantasies in prompting mentalizing, and of beneficent motivations in prompting non-mentalizing.

There can be cases, then, where MBT may primarily serve to further facilitate understanding of the thoughts and motivations of others, honing existing skills in relational aggression. A recent study by Gillespie, Kongerslev, Sharp, and colleagues has followed up on this concern. This work builds from Sharp's previous collaboration with Fonagy on hypermentalizing in adolescents with BPD. Gillespie and colleagues conducted a study with 80 male adolescents incarcerated mostly for violent offences. They assessed participants' ability to recognize emotions in others based on images of the eye region, intended as a measure of skills in mentalizing their display of feelings. Participants also completed questionnaire measures of aggression, psychopathic tendencies, and personality disorder. Success in identifying emotions from just the images of the eye region was associated with self-reported proactive aggression against others—but not reactive aggression—even controlling for psychopathic tendencies. The authors theorized that, when an individual has weak access to, and knowledge of, their own feelings, they are likely to also 'show problems in emotional resonance, that is, the ability to feel what another is feeling.'⁹² Despite previous reservations about the word (see Chapter 3), this capacity has recently been described by Bateman, Fonagy, and Campbell as 'empathy'.⁹³ Gillespie and colleagues argued that knowledge of the feelings of others can be used as a resource for relational aggression. The authors urged recognition that MBT may be actively harmful for patients with both conduct problems and psychopathic tendencies.⁹⁴ A similar concern has been raised about MBT by other commentators, such as Crittenden and Landini.⁹⁵

⁹² Gillespie, S. M., Kongerslev, M. T., Sharp, C., Bo, S., and Abu-Akel, A. M. (2018). 'Does Affective Theory of Mind Contribute to Proactive Aggression in Boys with Conduct Problems and Psychopathic Tendencies?' *Child Psychiatry & Human Development*, 49(6): 906–916, p. 912. cf. Davidsen, A. S. and Fosgerau, C. F. (2015). 'Grasping the Process of Implicit Mentalization'. *Theory & Psychology*, 25(4): 434–454.

⁹³ Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Borderline Personality Disorder', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 323–334, p. 337. For a valuable recent contribution to specifying the relationship between RF and empathy, see Borelli, J. L., Stern, J. A., Marvin, M. J., Smiley, P. A., Pettit, C., and Samudio, M. (2020). 'Reflective Functioning and Empathy among Mothers of School-Aged Children: Charting the Space between'. *Emotion*, Early View.

⁹⁴ The Gillespie paper, and other such efforts to address the potential harm of MBT with patients with histories of relational aggression, remain hindered by problems in the conceptualization of affect within accounts of mentalizing. In particular, it remains unclear whether affective mentalizing refers to efforts to mentalize feelings or the use of feelings in mentalizing. Additionally, the definitions of mentalizing considered in Chapter 4 would appear—perhaps inadvertently—to exclude the use of feelings from the definition of mentalizing except insofar as they help individuals deploy, conceive of, or reconsider mental states; the use of emotional resonance to make sense of mental states is, currently, technically outside of the definition of mentalizing. An exception is the encompassing definition from Fonagy, P. (1996). 'The Significance of the Development of Metacognitive Control over Mental Representations in Parenting and Infant Development'. *Journal of Clinical Psychoanalysis*, 5(1): 67–86: 'The psychological processes underpinning the view of oneself and others as motivated by mental states' (p. 74). However, what this definition gains in inclusiveness, it loses in precision, which is no doubt part of the reason it has been superseded over the decades.

⁹⁵ Crittenden, P. M. and Landini, A. (2011). *Assessing Adult Attachment: A Dynamic-Maturational Approach to Discourse Analysis*, New York: W. W. Norton & Company: 'we are wary of encouraging very disturbed individuals [to] use mentalization for self-protective functions because this may concurrently endanger others' (p. 39).

The wider social system

Introduction

In the previous chapter, we described developments in mentalization-based therapy (MBT), as well as a few of its current attributes and salient outstanding questions. One set of questions that was left for this chapter was the relationship between individual mentalizing—and mentalizing-focused intervention—and the individual's wider social and cultural ecology. This includes the matter of whether, in certain kinds of environment, improvement in mentalizing and/or epistemic trust may be harmful. On a wider scale, critics have alleged that, in attempting to adapt to the individual-centric environment of contemporary health provision, MBT itself has been complicit with the atomization of society.¹ Yet, in recent years, Fonagy and his collaborators have become increasingly critical of individual-centric modes of explanation, placing greater emphasis on the wider social system and the role of human interdependence in facilitating or hindering mentalization and mental health. These themes have been especially at the forefront recently during the 2020 and 2021 COVID-19 lockdowns.² Fonagy has advocated specifically that 'we must use the health catastrophe to drive social change'.³

The chapter begins by examining the growing attention to the social environment shown by Fonagy and colleagues, and especially their exploration of the role of friends and friendships for mentalization and epistemic trust. We will then examine the reflections and research by Fonagy and collaborators on public mental health. The researchers' hopes regarding school-based prevention will be given particular attention, and the chapter will also show how this work has shaped Fonagy's efforts as a policy influencer. Finally, the chapter will appraise the considerations offered by Fonagy and colleagues of the role of culture, in particular the issue of whether attention to cultural processes should be regarded as mentalizing, non-mentalizing or as not mentalizing, and whether organizations and societies can themselves be said to institutionalize cultures of mentalizing or non-mentalizing.

Reflecting on his own autobiography, Fonagy has acknowledged that he would likely have committed suicide as an adolescent had it not been for the intervention of his neighbour, who helped him receive treatment from Anne Hurry (see Chapter 2). Conversations with his colleagues Peter Fuggle and Dickon Bevington at the Anna Freud Centre have also had a profound role in highlighting to Fonagy the importance of the wider social system around the individual. Fuggle and Bevington pioneered Adaptive Mentalization-Based Integrative Treatment (AMBIT), an approach to the delivery of supportive services concerned to

¹ E.g. Rustin, M. (2015). 'Psychotherapy in a Neoliberal World'. *European Journal of Psychotherapy & Counselling*, 17(3): 225–239.

² See Allison, L. (2020). 'On Learning from Loss: Rereading 'Mourning and Melancholia'. Accessed at: <https://www.ucl.ac.uk/psychoanalysis/learning-loss-rereading-mourning-and-melancholia>; Fonagy, P. (2020). Tweet, 22 March. Accessed at: <https://twitter.com/peterfonagy/status/1241676565387382786?s=11>: 'Current crisis shows just how dependent we are on each other.'

³ Accessed at: <https://twitter.com/peterfonagy/status/1340576679077552129?s=11>

nurture the capacity of these services to sustain mentalizing and epistemic trust.⁴ The emphasis of Fuggle and Bevington on the organizational context of mentalizing has helped Fonagy to see that ‘mental health is a community responsibility’.⁵ In the 2019 edition of the *Handbook of Mentalising in Mental Health Practice*, Bateman and Fonagy have expressed their wish to ‘improve long-term outcomes in interpersonal function, a domain known to have only moderately good outcomes following MBT for BPD [borderline personality disorder]’.⁶ Elsewhere in the *Handbook of Mentalising in Mental Health Practice*, Fonagy and colleagues urged colleagues to recognize ‘the value of thinking about ways in which a social climate can be encouraged to become more mentalising’.⁷ And they closed their introduction to the volume by stating that ‘it is this wider social context of mentalising, and our human collective responsibility to support it, that will be at the core of mentalising endeavours in the future’.⁸ Bevington and Fuggle have observed a coming transformation in the very conceptualization of mentalising, because ‘mentalising is increasingly understood less as an individual power or even as a dyadic transaction, and more as a social and contextually determined activity’.⁹ This transition in conceptualizing mentalizing may be supported by the Covid pandemic and lockdowns, which have highlighted the fundamental contribution of threats and supports in the social environment to mentalizing capacities.¹⁰

The social environment

In individual therapy, the patient may learn to drop epistemic vigilance and consider the personal relevance of interactions with the therapist. Fonagy and colleagues perceive that doing so is relatively unlikely to be harmful, because the therapist’s intentions are largely benign and dependable, for all that the psychoanalytic unconscious may convey countervailing harmonics. The MBT therapist ‘consistently recognises the patient’s agency, focuses on the patient as an actor, and negotiates from the perspective of the patient’s self’.¹¹

⁴ Bevington, D. and Fuggle, P. (2019). ‘AMBIT: Engaging the Client and Community of Minds’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 211–228.

⁵ Kirby, T. (2019). ‘Peter Fonagy—battling the enemy of loneliness.’ *The Lancet Psychiatry*, 6(12): 987.

⁶ Bateman, A., Unruh, B. and Fonagy, P. (2019). ‘Individual Therapy Techniques’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 103–116, p. 108.

⁷ Fonagy, P., Campbell, C. and Allison, E. (2019). ‘Therapeutic Models’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 178.

⁸ Bateman, A. and Fonagy, P. (2019). ‘Introduction’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 3–20, p. 18. See also Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre: ‘When we consider the extensive evidence of the impact of parents’ and carers’ mental health on the mental health and wellbeing of children and young people and the influence of community and neighbourhood on both children and families, the narrow focus of most treatment approaches on individual dysfunction is surprising.’

⁹ Bevington, D. and Fuggle, P. (2019). ‘AMBIT: Engaging the Client and Community of Minds’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 211–228, p. 221. See also Jurist, E. and Sosa, M. P. (2019). ‘Commentary on Mentalization and Culture’. *Clinical Psychology: Science and Practice*, 26(4): e12302.

¹⁰ Lassri, D. and Desatnik, A. (2020). ‘Losing and regaining reflective functioning in the times of COVID-19: Clinical risks and opportunities from a mentalizing approach’. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1): S38.

¹¹ Fonagy, P., Campbell, C. and Allison, E. (2019). ‘Therapeutic Models’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 175.

In group therapy, this is more challenging as the wishes and plans of others may not be benign or dependable, and individuals may feel unsafe with the group. This has been reported as the most serious concern of patients undergoing MBT, especially those who attended groups with rolling membership.¹² Bateman and Fonagy have presumed that, if well handled by the therapist, vulnerability and willingness to learn in the group context should not be harmful. In fact, they have argued that overcoming the challenges to mentalizing and epistemic trust in this potentially turbulent context will have additional benefit beyond that available from individual therapy: ‘the power of group therapy [lies in its potential] to stimulate the capacity of the patient to manage anxiety within highly charged circumstances while maintaining mentalising. It is in the group that patients can truly balance emotional states evoked in a complex situation and their ability to continue mentalising.’¹³ However, the difference in efficacy between individual and individual + group MBT has yet to be tested, and some patients report that they regard the group therapy as lacking benefit for them.¹⁴ Inderhaug and Karterud, and Folmo and colleagues, have also cautioned that Bateman and Fonagy may be over-optimistic about the capacity of the group therapist to achieve a safe and constructive group atmosphere.¹⁵ Karterud has reported findings that patients who began MBT with low reflective functioning scores on the Adult Attachment Interview responded especially well to individual therapy and especially poorly to group-based therapy.¹⁶

In contrast to both individual and (perhaps) group therapy, in which a certain safety can be ensured, Bateman and Fonagy acknowledged that the patient’s wider life outside therapy may well be characterized by danger and chaos—and is very likely to remain so at the end of MBT.¹⁷ This may include interactions with state services in which professionals appear far from benign or dependable.¹⁸ Allison has described how this may play out in the case of homelessness:

Many people who have chronic or repeat experiences of homelessness report traumatic childhood experiences such as abuse, neglect, bullying, witnessing alcoholism, or domestic violence. These difficult beginnings can leave them profoundly mistrustful of others. Not surprisingly, then, they are also often mistrustful of services and may refuse to engage or struggle to remain engaged, especially if the help on offer comes with rigid conditions

¹² Lonergáin, D. Ó., Hodge, S., and Line, R. (2017). ‘Service User Experiences of Mentalisation-Based Treatment for Borderline Personality Disorder’. *Mental Health Review Journal*, 22(1): 16–27.

¹³ E.g. Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 156.

¹⁴ Lonergáin, D. Ó., Hodge, S., and Line, R. (2017). ‘Service User Experiences of Mentalisation-Based Treatment for Borderline Personality Disorder’. *Mental Health Review Journal*, 22(1): 16–27: ‘Most participants described wanting individual sessions only and some did not see the purpose of group MBT’ (p. 19).

¹⁵ Inderhaug, T. S. and Karterud, S. (2015). ‘A Qualitative Study of a Mentalization-Based Group for Borderline Patients’. *Group Analysis*, 48(2): 150–163. See also Folmo, E. J., Karterud, S. W., Kongerslev, M. T., Kvarstein, E. H., and Stånicke, E. (2019). ‘Battles of the Comfort Zone: Modelling Therapeutic Strategy, Alliance, and Epistemic Trust—A Qualitative Study of Mentalization-Based Therapy for Borderline Personality Disorder’. *Journal of Contemporary Psychotherapy*, 49(3): 141–151.

¹⁶ Karterud, S. (2011). ‘Constructing and Mentalizing the Matrix’. *Group Analysis*, 44(4): 357–373, p. 383.

¹⁷ Allison and Campbell warn against any assumption that socio-economic adversity will necessarily entail benefits to epistemic distrust. Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre: ‘it is important not to make simplistic assumptions about the quality of relationships implied by socio economic disadvantage. There is evidence that individuals who are less socio economically privileged tend to behave in more community and socially oriented ways in interpersonal trust experiments than more affluent individuals (Dubois, Rucker, & Galinsky, 2015)’, citing Dubois, D., Rucker, D. D., and Galinsky, A. D. (2015). ‘Social Class, Power, and Selfishness: When and Why Upper and Lower Class Individuals Behave Unethically’. *Journal of Personality and Social Psychology*, 108(3): 436–449.

¹⁸ Mason, C., Taggart, D., and Broadhurst, K. (2020). ‘Parental Non-Engagement within Child Protection Services—How Can Understandings of Complex Trauma and Epistemic Trust Help?’. *Societies*, 10(4): 93.

attached. When they do try to get help and the attempt backfires, this further undermines their trust in the people around them.¹⁹

In fact, the interaction between current adversities and mentalizing difficulties may be sufficiently intense that this intrudes on psychotherapy. Therapists may be drawn in to the non-mentalizing dynamics of the wider social system, even as they try their best to help: 'the practitioner becomes embedded within the client's social survival mechanism and this survival mechanism tends to work to destroy balanced mentalizing'.²⁰

Fonagy and colleagues fully accept that in an environment where others' wishes and plans are not benign or dependable, epistemic vigilance is entirely appropriate.²¹ When therapy leads epistemic vigilance to be lowered without discrimination, a deterioration in the patient's mental health can be expected.²² The idea that a patient's capacity and willingness to learn from experience is responsive to the qualities and dependability of their social environment had been a long-standing tenant of the Anna Freudian tradition for child psychoanalysis (see Chapter 7). It was anticipated that psychoanalysts would work collaboratively with parents to make the child's home environment feel secure and safe enough to facilitate the adoption by the child of new forms of adaptation.²³ As a rule, children are generally more responsive to social environments than adults.²⁴ Nonetheless, Fonagy and colleagues have increasingly perceived that for adults, too, the environment needs to be secure and safe enough for epistemic trust to be appropriate.

This was, in fact, already recognized by Fonagy back in 2003: 'individual therapy, while prototypical of the process of rekindling mentalization, may not cut it on its own. It is in the shared communally constructed creation of mind, in the patient-therapist dyad, in the family, in the psychotherapy group, in the multidisciplinary team, or in the organization of the entire treatment'.²⁵ However, in the early 2000s, Fonagy had no conceptual model available for thinking about the communally constructed creation of mind, so this point was not elaborated. In recent years, though, the concept of epistemic trust has been drawn upon for considering how an individual can benefit from their communities. Fonagy and colleagues

¹⁹ Allison, E. (2015). 'Epistemic Trust: A New Perspective on the Barriers to Change in Chronic and Repeat Homelessness', in *What Works: Rethinking Hopelessness*, Cobham, UK: Berkeley Foundation, p. 9. Accessed at: https://www.berkeleyfoundation.org.uk/media/pdf/r/r/Rethinking_Homelessness.pdf. Some analogues can be drawn, in some cases, with the experiences of state services among refugees. See the remarks by Fonagy in Novikova, M. (2019). 'Can a Form of Therapy Focused on Understanding the Mental State of Others Help Refugees in the UK?', *Guardian*, 21 June. Accessed at: <https://www.theguardian.com/the-guardian-foundation/2019/jun/21/how-to-make-integration-mutual>.

²⁰ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

²¹ See also Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., and Wilson, D. (2010). 'Epistemic Vigilance'. *Mind & Language*, 25(4): 359–393.

²² Allison, E. and Fonagy, P. (2016). 'When is Truth Relevant?'. *Psychoanalytic Quarterly*, 85(2): 275–303: 'If the truth that the lifting of epistemic hypervigilance uncovers is unremitting hostility and the absence of benign influence, the recovery of epistemic trust through therapy will generate no lasting improvement and may even lead to deterioration' (p. 295).

²³ Fonagy, P., Bleiberg, E., and Target, M. (1997). 'Child Psychoanalysis: Critical Overview and a Proposed Reconsideration'. *Child and Adolescent Psychiatric Clinics of North America*, 6: 1–38.

²⁴ Fonagy, P. and Sharp, C. (2008). 'Treatment Outcome of Childhood Disorders: The Perspective of Social Cognition', in C. Sharp, P. Fonagy, and I. Goodyer (eds), *Social Cognition and Developmental Psychopathology*, New York: Oxford University Press, pp. 411–470: 'Interventions in young children draw their effectiveness from the parent's influence on social cognition, while in middle childhood a combination of problem-solving and parent training might be most effective. By adolescence, social cognitive patterns are so well established that only a combination of a number of systems is likely to shift the young person's social cognitive stance' (p. 446).

²⁵ Fonagy, P. (2003). 'Clinical Implications of Attachment and Mentalization: Efforts to Preserve the Mind in Contemporary treatment. Epilogue'. *Bulletin of the Menninger Clinic*, 67: 271–280, p. 277.

have claimed that epistemic trust will have most benefit for patients if it is accompanied by the capacity to form and make use of stable relationships beyond therapy. And vice versa. Allison and Fonagy have concluded that, in successful MBT, it is the 'evolutionary capacity for learning from a social situation that is rekindled', and within this, 'learning who one can be friends with is key'.²⁶ Or, as Fonagy, Luyten, Allison, and Campbell have argued: 'Enhanced mentalizing allows the patient to achieve improved social relationships and recognize who is a reliable and trustworthy source of information—that is, who one can “be friends with”'.²⁷ Signalling its importance, this phrase is repeated verbatim in the *Handbook of Mentalising in Mental Health Practice*.²⁸

The concept of the 'friend' is a long-standing, if subtextual, concern in Fonagy's writings. As we saw in Chapter 1, he has repeatedly recalled the importance of a lack of any friends among the conditions that contributed to his feelings of depression and suicidality as an adolescent, and that led him into therapy with Anne Hurry. The theme of the 'friend' in Fonagy's academic writings first appeared in his 1992 tribute to George Moran, where he described how much of both personal and professional relevance he had gained from the friendship.²⁹ The friendship was deeply felt: when Moran was dying, Fonagy took sabbatical leave from work to provide care and support.³⁰ On assuming the position of director of the Anna Freud Centre, Fonagy emphasized the influence on his work of his 'very special friendship' with Moran.³¹ Fonagy's descriptions of the intellectual aspects of the friendship resonate with the ideas of the anthropologist Ed Hutchins regarding 'distributed cognition' (see Chapter 7), in which certain thoughts or insights are possible thanks to the properties emergent within an interaction, rather than solely as the sum of the cognitive properties of the members of that interaction. For Fonagy, his weekly conversations with Moran about their clinical work, which they called the 'Saturday club', provided scaffolding for reflection.³² The conversations brought together the tradition of Anna Freudian psychoanalysis, the problem of insulin control of diabetic patients, the expectation of regular future conversation—among other factors—to form a semi-autonomous ecology for supporting thinking that could not have been achieved individually.

In Fonagy's subsequent thinking, friendships were treated as something rather special. Even when other relationships lack the potential for playfulness, or are hindered by conflict or forms of non-mentalizing interaction, friendships carve out space for their own semi-autonomous functioning.³³ Indeed, Fonagy has collaborated on innovative longitudinal

²⁶ Allison, E. and Fonagy, P. (2016). 'When is Truth Relevant?'. *Psychoanalytic Quarterly*, 85(2): 275–303, pp. 295–296.

²⁷ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9. Bateman and Fonagy have also proposed that 'perhaps the most important factor' holding back temporary failures of mentalizing from doing harm are 'the normally corrective response of the other person with whom one is interacting'. Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 125.

²⁸ Fonagy, P., Campbell, C. and Allison, E. (2019). 'Therapeutic Models', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 176.

²⁹ Fonagy, P., Moran, G. S., and Target, M. (1992). 'Aggression and the Psychological Self'. *Bulletin of the Anna Freud Centre*, 15: 269–284, p. 269.

³⁰ Personal communication, Peter Fonagy, December 2019.

³¹ Fonagy, P. (2003). 'The Anna Freud Centre: About the Directors'. Originally at <https://www.annafreudcentre.org/fonagy.htm>. Accessed at: <https://web.archive.org/web/20060925220711/>.

³² Ezrati, O. (2014). 'Freud Off: Giving New Meaning to Psychoanalysis'. Haaretz, 5 April. Accessed at: <https://www.haaretz.com/life/books/.premium-giving-new-meaning-to-psychoanalysis-1.5243899>.

³³ A few years later, in 1997, Fonagy highlighted Dunn's findings that 'individual differences found in pretend play, management of conflict, and talking about mental states are not correlated between social situations

work led by van Harmelen examining longitudinal predictors of the degree to which an individual functions better or worse than expected in adulthood, given their early childhood family experiences. Van Harmelen and colleagues found that, in the context of adverse experiences of early care, positive adolescent friendships predict resilience in adulthood better than do positive family relationships in adolescence.³⁴ Fonagy has elsewhere also acknowledged that the capacity to trust in friendships in adolescence and afterwards is certainly shaped by family relationships.³⁵ Nonetheless, for Fonagy, the true friend is someone who creates a social environment within which epistemic vigilance can be dropped, even if epistemic vigilance must be retained in other contexts. This can entail shared non-mentalizing, as in shared pretend mode with its attendant potential benefits and problems. However, the fall of epistemic vigilance can also make possible access to the minds of others as a source of personal learning and development, in acknowledgement of the limits of our current understanding of ourselves and the world.³⁶

Fonagy's awareness of the debt we can have to friends for learning and development, as well as for remaining mentally well, has perhaps shaped his unusual diligence in writing detailed, loving obituaries and tributes for late friends and colleagues over his career. It is difficult to think of any other major research psychologist who, across their career and not just in retirement, has written so many obituaries. Fonagy has also frequently described the benefit and pleasure he has gained from thinking and writing with academic friends in pursuing collaborative work.³⁷ At a more general level, Fonagy and colleagues have repeatedly stated, based on their experiences, that a clinician's capacity to mentalize is sustained or undermined by the support they receive to do so from supervision, colleagues, and from wider

(mother, siblings, close friend)'. Fonagy, P. and Target, M. (1997). 'Attachment and Reflective Function: Their Role in Self-Organization'. *Development and Psychopathology*, 9(4): 679–700, p. 689. The implication is that the relational supports for mentalizing are relatively independent across these different relational forms. See also Cheftetz, R. A. (2013). 'A Fluctuating Capacity to Mentalize: Affect Scripts and Self-State Systems as (not so) "Strange Attractors": A Discussion of Margy Sperry's "Putting our Heads Together: Mentalizing Systems"'. *Psychoanalytic Dialogues*, 23(6): 708–714.

³⁴ Van Harmelen, A. L., Kievit, R. A., Ioannidis, K., Neufeld, S., Jones, P. B., Bullmore, E., ... and NSPN Consortium. (2017). 'Adolescent Friendships Predict Later Resilient Functioning across Psychosocial Domains in a Healthy Community Cohort'. *Psychological Medicine*, 47(13): 2312–2322.

³⁵ E.g. Fonagy, P. (2019). 'The Future Prospects of Mentalization Based Therapies', 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November: 'Close family networks are maximally beneficial in childhood but only if they later transition to more diverse varied (large) or family and friend (mixed) networks. Supportive family relationships in childhood provide a secure base from which to later diversify social networks', citing Manalel, J. A. and Antonucci, T. C. (2020). 'Beyond the Nuclear Family: Children's Social Networks and Depressive Symptomology'. *Child Development*, 91(4): 1302–1316. See also Fonagy, P. (1999). 'The Transgenerational Transmission of Holocaust Trauma'. *Attachment & Human Development*, 1(1): 92–114: 'Glen's material contained many ideas from the Holocaust; he appeared to experience these as his own, notwithstanding the distance of two generations. We have noted the profound impact of the film *Schindler's List* on his fantasies, as well as his paranoid anxieties of persecution. Perhaps even more relevant, his thinking was permeated by specific images which we can trace to his grandmother's experiences, as probably imagined by his mother. In this category I would include: his cruel work regime, used to obliterate psychic reality; his terror of being mocked or humiliated by those he considered his friends' (p. 105).

³⁶ There are significant alignments between Fonagy's discussion of the figure of the friend and the proposals of Gadamer. See Gadamer, H. G. ([1985] 1999). 'Friendship and Self-Knowledge: Reflections on the Role of Friendship in Greek Ethics', in *Hermeneutics, Religion, and Ethics*, trans. Joel Weinsheimer, New Haven, CT: Yale University Press.

³⁷ E.g. Fonagy, P. and Target, M. (2006). 'The Mentalization-Focused Approach to Self Pathology'. *Journal of Personality Disorders*, 20: 544–576: 'The work summarised in this paper is the result of a collaborative effort of a group of wonderfully talented individuals who have honoured the author with their friendship over the past years.' George Moran, György Gergely, Miriam and Howard Steele, Helen Stein, John Allen, Efrain Bleiberg, Anthony Bateman, and Liz Allison are listed (p. 544).

institution systems and their organizational cultures.³⁸ The clinician needs support if he or she is to face down the 'hours of sometimes immensely painful analytic listening, toleration of distress and uncertainty, and moments of almost indescribable emptiness'.³⁹

Fonagy and colleagues have highlighted the contribution of friends to emotion regulation and mentalizing, for good or for ill. For instance, Fonagy and Luyten have drawn attention to a prospective study by Crick and colleagues that showed that 'the best predictors of BPD features from fourth to sixth grade were indicators of social dysfunction such as friend exclusivity (overly close relationship with friend), relational aggression (impulsivity), and cognitive sensitivity (hostile, untrusting paranoid world view)'.⁴⁰ Fonagy and Luyten theorized that difficulties in social relationships can set off cascades of contributors to mental health difficulties, such as reciprocal relationships between non-mentalizing and difficulties with affect modulation.⁴¹ By contrast, more positively, Twemlow, Fonagy, and colleagues were interested that their mentalization-based intervention in schools (see below) had a very large effect on children's reports of being able to form and make use of school-based friendships. The number of children who said that they found making friends easy doubled following the intervention, from 30% to 60% of pupils; the number willing and able to make friends with pupils of a different race increased from 30% to 50%.⁴² Such findings suggest important links between friendships and epistemic trust. Yet a concern with these two matters is not currently a priority in many therapeutic modalities. For instance, neither is a special priority within child psychoanalysis. In the Anna Freud Centre retrospective study follow-up, Fonagy and Target found that former patients at the Centre reported *more* difficulties forming and making use of friendships than their non-treated siblings.⁴³ Fonagy and colleagues have also criticized the excessive focus on clinical indicators in evaluating the effectiveness of mental health interventions, to the neglect of day-to-day functioning and the capacity to form supportive friendships.⁴⁴ This is contrary to research on what patients themselves report wanting from therapy.⁴⁵

³⁸ E.g. Bevington, D., Fuggle, P., Cracknell, L. and Fonagy, P. (2017). *Adaptive Mentalisation-Based Integrative Treatment: A Guide for Teams to Develop Systems of Care*, Oxford: Oxford University Press: 'Helping interventions (of whatever orientation or model) are situated within at least three interconnected systems: the family/care setting, the intervening team, and the wider welfare and educational network. For the worker, one key challenge is to continue working effectively when the interconnected systems create unintended contradictions or negative feedback patterns (what we describe as "dis-integration")' (p. 23).

³⁹ Fonagy, P. (1996). 'Commentaries'. *Journal of the American Psychoanalytic Association*, 44: 404–422, p. 408.

⁴⁰ Fonagy, P. and Luyten, P. (2009). 'A Developmental, Mentalization-Based Approach to the Understanding and Treatment of Borderline Personality Disorder'. *Development and Psychopathology*, 21(4): 1355–1381, p. 1356. Discussing Crick, N. R., Murray-Close, D., and Woods, K. (2005). 'Borderline Personality Features in Childhood: A Short-Term Longitudinal Study'. *Development and Psychopathology*, 17(4): 1051–1070.

⁴¹ Luyten, P., Fonagy, P., Lemma, A., and Target, M. (2012). 'Depression', in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 385–418, p. 402.

⁴² Twemlow, S. W., Fonagy, P., Sacco, F. C., Vernberg, E., and Malcom, J. M. (2011). 'Reducing Violence and Prejudice in a Jamaican All Age School Using Attachment and Mentalization Theory'. *Psychoanalytic Psychology*, 28: 497–511, p. 508.

⁴³ Fonagy, P. and Target, M. (2002). 'The History and Current Status of Outcome Research at the Anna Freud Centre'. *The Psychoanalytic Study of The Child*, 57(1): 27–60, p. 48.

⁴⁴ E.g. Barlow, J., Barnes, J., Sylva, K., Fonagy, P., and Fearon, P. (2016). 'Questioning the Outcome of the Building Blocks Trial'. *Lancet*, 387(10028): 1615–1616.

⁴⁵ E.g. Katsakou, C., Marougka, S., Barnicot, K., Savill, M., White, H., Lockwood, K., and Priebe, S. (2012). 'Recovery in Borderline Personality Disorder (BPD): A Qualitative Study of Service Users' Perspectives'. *PloS One*, 7(5): e36517; Katsakou, C. and Pistrang, N. (2018). 'Clients' Experiences of Treatment and Recovery in Borderline Personality Disorder: A Meta-Synthesis of Qualitative Studies'. *Psychotherapy Research*, 28(6): 940–957. See also Fonagy, P. (2010). 'The Changing Shape of Clinical Practice: Driven by Science or by Pragmatics?'. *Psychoanalytic Psychotherapy*, 24(1): 22–43: 'From a professional's standpoint, as from that of the ordinary member of the public, physical role limitation, physical function and pain have high priority, while those suffering disorders rate dignity and general wellbeing (mood, global assessment of life, having a partner, job, lots of social contact) as more

Fonagy has stated that he has often reflected on what constitutes the difference between a professional and a friend, reflections that appear to have influenced the injunction in the manual that the therapist should avoid posturing in ways that signal their expertise.⁴⁶ Bateman and Fonagy have argued that an MBT therapist should ‘first engage in the process of “being ordinary” ... If in doubt, say to the patient what you would say to a good friend if he/she was telling you the same story while sitting in a cafe and you wanted to transmit a sense that you were “getting” their emotional state.’⁴⁷ The relationship with the therapist models the ordinary opportunities friendships provide for recognizing ostensive cues, and for gaining skills and confidence in deploying mentalizing and epistemic trust:

Meaningful change is thus possible only if the person can use their social environment in a positive way (and if the social environment is sufficiently supportive to allow this to happen). For this to happen, recognition of self-agency is key, and this recognition is best achieved through the ostensive cues that are provided by feeling appropriately mentalised by another person.⁴⁸

However, the therapeutic context is not only a model for forming and making good use of friendships. It also provides a space in which the patient can reflect on and appraise their social interactions. It seeks to help the patient consider how to use mentalization and drop epistemic vigilance in discriminate ways, appropriate to the opportunities and threats of the environment.⁴⁹ The most effective treatment, Fonagy has argued, is one that ‘enables the individual to access and support, find inspiration and consolation, from a whole range of other people.’⁵⁰ However, both inspiration and consolation entail the discrimination of appropriate occasions for mentalization and epistemic trust.

In the 2016 MBT manual, Bateman and Fonagy encouraged therapists to act as an advocate for the patient with other services: ‘MBT recommends that early in therapy the patient’s social context is stabilized. Change will be impossible if housing, financial, employment,

important. Wellbeing should feature at least alongside, if not in place of, lists of symptoms in outcome studies (Pressman and Cohen, 2005)’ (p. 34).

⁴⁶ Fonagy, P. (1998). ‘Prevention, the Appropriate Target of Infant Psychotherapy’. *Infant Mental Health Journal*, 19(2): 124–150: ‘On a personal note, I have to confess to having often felt uneasy about the mixture of roles and moralities involved in a psychotherapeutic relationship (friend and helper on the one hand, professional and expert on the other)’ (p. 131).

⁴⁷ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 242. This prioritization of ordinariness follows very much along the same lines as Bion, who would tell patients at the start of therapy, ‘this is just an ordinary conversation about a very delimited range of experience—namely your emotional experience—and what happens to you if you cannot have emotional experience’. Culbert-Koehn, J. (2011). ‘An Analysis with Bion: An Interview with James Gooch’. *Journal of Analytical Psychology*, 56(1): 76–91, p. 82. On the ‘ordinary’ as a rich and complex resource for interaction, see Cavell, S. (1994). *In Quest of the Ordinary: Lines of Skepticism and Romanticism*, Chicago: University of Chicago Press.

⁴⁸ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 35.

⁴⁹ Ibid.: ‘We suggest that successful treatments all involve three essential systems of communication relating to epistemic trust ... Communication system 1: communication of therapeutic model-based content ... Serves as an ostensive cue that increases the patient’s epistemic trust and thus acts as a catalyst for therapeutic success (“therapeutic alliance by any other name”). Communication system 2: mentalising as a common factor: the therapeutic setting serves to increase the patient’s mentalising. Communication system 3: social learning in the context of epistemic trust: The patient applies his/her restored mentalising in the wider (social) environment, which reinforces and builds upon what he/she has learned in therapy’ (p. 28).

⁵⁰ Fonagy, P. (2016). ‘We have Hard Choices to Make on Children’s Mental Health’. Huffington Post, 10 October. Accessed at: http://www.huffingtonpost.co.uk/peter-fonagy/world-mental-health-day_b_12429138.html.

probation and other stresses are dominant. The MBT clinician is an active advocate for the patient's link to the wider social system.⁵¹ There seems a clear turn in the past few years in accounts of MBT towards depictions of an active role of the clinician in supporting the patient to end harmful relationships, build on and nurture friendships that appear more beneficial, stabilize their living conditions, and other strategies to improve the social environment outside of therapy.⁵² Fonagy and Allison have put forward the empirical prediction that 'psychotherapy for BPD is much more likely to succeed if the individual's social environment at the time of treatment is largely benign. Although we do not know of any systematic studies that have explored this moderator.'⁵³ They have urged further work to address this question.⁵⁴

Prevention and public mental health

One way in which the importance of social context has been registered in the work of Fonagy and colleagues over the years has been in their considerations of prevention and public mental health. The role of poverty, discrimination, and other social adversities and forms of oppression have long been registered as powerful contributors to mental illness.⁵⁵ However, through the 1990s and early 2000s, the primary answer Fonagy and collaborators offered to these factors was to situate individual regulatory capacities as potential buffers for the effects of adversity. This reflected and responded to the growing focus on individual self-management, rather than collective solutions to social problems, in culture and policy more generally in these decades (Chapter 2). In their 2002 paper, 'Early Intervention and the Development of Self-Regulation', Fonagy and Target offered their view that there were no 'effective models of intervention' to address these society-wide factors. In the context of apparent fatalism about social change, they instead advocated a 'move toward person variables that were of particular importance in making an individual vulnerable to risk'. A survey of such potential variables led them to conclude that 'self-regulation is currently the most promising candidate'.⁵⁶ Fonagy and Higgitt would later re-state the point, emphasizing the importance of childhood as a period in which self-regulation capacities are developed:

It does not necessarily follow that social causes have social solutions; aspirin will relieve a headache even if the cause is poverty. Thus, it is likely that an understanding of the

⁵¹ Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, p. 34.

⁵² E.g. Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Antisocial Personality Disorder in Community and Prison Settings', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 335–349.

⁵³ Fonagy, P. and Allison, E. (2014). 'The Role of Mentalizing and Epistemic Trust in the Therapeutic Relationship'. *Psychotherapy*, 51(3): 372–380, p. 379.

⁵⁴ One step in this direction was undertaken by Bateman and Fonagy who developed a mentalisation-based intervention for families of people with BPD. Bateman, A. and Fonagy, P. (2019). 'A Randomized Controlled Trial of a Mentalization-Based Intervention (MBT-FACTS) for Families of People with Borderline Personality Disorder'. *Personality Disorders*, 10(1): 70–79.

⁵⁵ E.g. Fonagy, P., and Higgitt, A. (2000). 'Early Influences on Development and Social Inequalities: An Attachment Theory Perspective', in A. R. Tarlov and R. F. S. Peter (eds.), *The Society and Population Health Reader*, Volume 2: A state and community perspective, New York: New Press, pp. 104–130; Murphy, M. and Fonagy, P. (2013). 'Mental Health Problems in Children and Young People', in Department of Health, *Our Children Deserve Better, Prevention Pays: Annual Report of the Chief Medical Officer* (Chapter 10), London: Department of Health.

⁵⁶ Fonagy, P. and Target, M. (2002). 'Early Intervention and the Development of Self-Regulation'. *Psychoanalytic Inquiry*, 22(3): 307–335, pp. 312–313.

pathways of health inequalities will offer the possibility of policies that might directly contribute to addressing social inequalities in health. Social engineering is unlikely to be a viable way to address the social gradient problem. The essence of prevention is to find a point along a causal path that allows the possibility of cost-effective psychosocial manipulation and intervention. This implies intervention in childhood and finding those who are 'hard to reach'.⁵⁷

In a pair of papers from 2004, Fonagy argued that public resources are limited, and so need to be targeted carefully. He argued in favour of directing resources towards prevention, especially towards support for families with young children—for instance, through intensive prenatal mental health screening.⁵⁸ When prevention fails, he proposed that services should prioritize those with biomarkers and family history suggestive of risk of mental health issues. This would include even the use of biomarkers rather than children's behaviour in screening for child abuse, and the targeting of interventions for child maltreatment to those children with biomarkers indicative of greater responsiveness to intervention.⁵⁹ Yet, over the past 15 years, it would appear that Fonagy's stance has shifted. He has argued against his earlier stance that resources, and even child protection services, should be targeted on the basis of biomarkers and family history, perhaps due to reservations about the moral implications of such a position. Instead, he has proposed that differences in biomarkers and family history—much like differences in symptom profiles—may indicate groups who would benefit from different kinds of intervention when they show poor mental health and seek help.⁶⁰

Fonagy has also qualified his earlier rather exclusive focus on individual self-regulation as the point that permits cost-effective psychosocial manipulation and intervention. He has stated in interview his view that 'We should be looking at how we can work toward a more equal distribution of wealth, and in some ways try to prevent the extant inequalities from affecting future generations'⁶¹—but he has not advocated this position loudly or prominently. There is a sense in which Fonagy appears to remain pessimistic about the possibilities of social change, and of the evidence base for society-level interventions. This may be linked to the realist thread in his theorizing, which sees little value in optimism that change will be successful in the absence of supporting evidence and a practical delineation of intervention

⁵⁷ Fonagy, P. and Higgitt, A. (2007). 'The Early Social and Emotional Determinants of Inequalities in Health', in G. Baruch, P. Fonagy, and D. Robins (eds), *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, Chichester, UK: John Wiley & Sons, pp. 3–34, p. 4.

⁵⁸ Fonagy, P. and Higgitt, A. (2004). 'Early Mental Health Intervention and Prevention: The Implications for Government and the Wider Community', in B. Sklarew, S. W. Twemlow, and S. M. Wilkinson (eds), *Analysts in the Trenches: Streets, Schools, War Zones*, Mahwah, NJ: Analytic Press, pp. 257–309.: 'Prenatal screening should include a prenatal assessment of mental health so that appropriate psychological or medical interventions can begin at this time' (p. 267).

⁵⁹ Fonagy, P. (2004). 'Psychotherapy Meets Neuroscience: A More Focused Future for Psychotherapy Research'. *The Psychiatrist*, 28(10): 357–359: 'Knowing that in individuals with the S/S genotype severe maltreatment doubles the probability of major depressive disorder (to over 60% from 30% for those with the genotype) helps us to focus interventions on childhood maltreatment for the first group to a greater extent than for the L/L group' (p. 358). Fonagy, P. (2007). 'My Brain Mapper'. Accessed at: <https://www.ucl.ac.uk/news/2007/may/ucl-news-my-brain-mapper>: 'By using brain scans, rather than studying how children behave, we will be able to spot those who are in greatest need of intervention.'

⁶⁰ Fonagy, P. (2010). 'The Changing Shape of Clinical Practice: Driven by Science or by Pragmatics?'. *Psychoanalytic Psychotherapy*, 24(1): 22–43: 'The moral here is not that psychotherapy should not be offered to people without this or that allele, but rather that the mechanism by which therapy achieves its effect may be quite different for these constitutionally distinguishable groups of individuals' (pp. 37–38). See also Bakermans-Kranenburg, M. J., and van Ijzendoorn, M. H. (2011). 'Differential Susceptibility to Rearing Environment Depending on Dopamine-Related Genes: New Evidence and a Meta-Analysis'. *Development and Psychopathology*, 23(1): 39–52.

⁶¹ Jurist, E. L. (2010). 'Elliot Jurist Interviews Peter Fonagy'. *Psychoanalytic Psychology*, 27(1): 2–7: p. 5.

mechanisms (see Chapter 1). It may reflect Fonagy's decades of work as a psychoanalyst. Fonagy has himself acknowledged that the consulting room context of individual analysis can contribute to making individual problems and individual solutions most salient.⁶² The consulting room is a sequestered space a) in which the therapist and client can be concerned with mental states, b) in which the practical fall-out of mistakes about mental states is constrained and not fatal for the interaction, and c) such mistakes can themselves be attended to. Originally psychoanalysis, and subsequently MBT, have been facilitated by the consulting room as this sequestered space.⁶³ Indeed, it facilitates the claim that mentalization is the active ingredient in all good therapy, and as 'common ground' between modalities.⁶⁴ However, the sequestration also has consequences for what thoughts or feelings seem salient.

Fonagy's continued pessimism regarding social change may also, following Bollas, be regarded as influenced by the wider societal focus on individual responsibility for thoughts and feelings, which to a certain extent Fonagy's work reflects and to which it responds (see Chapter 2). It should be emphasized, however, that there is no intrinsic link between the theory of mentalization and a focus on individual solutions to social problems or fatalism regarding social change. For instance, Dickon Bevington, medical director at the Anna Freud Centre, is politically outspoken on issues of social justice.⁶⁵

By contrast, an increasing attention to the social context has led Fonagy to seek to examine how 'socially excluded groups such as LGBT individuals, asylum seekers, and those from poorer backgrounds' might be supported in their individual self-regulatory efforts.⁶⁶

In response to the murder of George Floyd by a Minneapolis police officer, and subsequent civil protests, Fonagy recently publicly reflected on his stance. In the past, he has generally

⁶² Interview with Peter Fonagy, cited in Maddox, L. (2018). *Blueprint: How Our Childhood Makes Us Who We Are*, London: Robinson, Chapter 7. See also Lee, N. N. (2014). 'Sublimated or Castrated Psychoanalysis? Adorno's Critique of the Revisionist Psychoanalysis: An Introduction to "The Revisionist Psychoanalysis"'. *Philosophy & Social Criticism*, 40(3): 309–338. There have been schools of therapy that give comparatively more weight to social and political context. However, the Anna Freudian tradition within which Fonagy trained was not one of them. Cf. Parker, I. (2011). *Lacanian Psychoanalysis: Revolutions in Subjectivity*, London: Routledge; Chancer, L. and Andrews, J. (2013). *The Unhappy Divorce of Sociology and Psychoanalysis: Diverse Perspectives on the Psychosocial*, London: Palgrave; Frosh, S. (2014). 'Psychoanalysis as Political Psychology', in P. Nesbitt-Larking, C. Kinnvall, and T. Capelos with H. Dekker (eds), *The Palgrave Handbook of Political Psychology*, Basingstoke: Palgrave, pp. 55–71; Cushman, P. (2015). 'Relational Psychoanalysis as Political Resistance'. *Contemporary Psychoanalysis*, 51(3): 423–459.

⁶³ Shapiro, S. A. (1996). 'The Embodied Analyst in the Victorian Consulting Room'. *Gender and Psychoanalysis*, 1(3): 297–322; Callard, F. (2014). 'Consulting Rooms: Notes towards a Historical Geography of the Psychoanalytic Setting', in P. Kingsbury and S. Pile, (eds), *Psychoanalytic Geographies*, Farnham, Ashgate, pp. 73–88. A difference between psychoanalysis and MBT here is materially signified by the couch, which helps focus the setting on the mental states of the patient, rather than the potential for negotiated identification of mental states and mistakes in both parties, as in MBT. Cf. Skolnick, N. J. (2015). 'Rethinking the Use of the Couch: A Relational Perspective'. *Contemporary Psychoanalysis*, 51(4): 624–648.

⁶⁴ Munich, R. L. (2006). 'Integrating Mentalization-Based Treatment and Traditional Psychotherapy to Cultivate Common Ground and Promote Agency', in J. Allen and P. Fonagy (eds), *Handbook of Mentalization-Based Therapy*, London: Wiley, pp. 143–156.

⁶⁵ It may also be noted that the 'On My Mind' resources on the Anna Freud Centre website include guidance for young people on activism as a form of individual self-care. Accessed at: <https://www.annafreud.org/on-my-mind/self-care/activism>: 'Not only does activism give you distraction from your own life, but it also allows you to stop others being distracted from important issues. You get to give back to a cause that you are passionate about, whilst also gaining self-gratification in return. Taking part in group campaigns can also be a great way of meeting others who share your views.'

⁶⁶ Fonagy, P., Hayes, D. and Stapley, E. (2019). 'Understanding and Facilitating Self-Management in Child and Youth Mental Health for Socially Excluded Populations'. Accessed at: https://clahrc-norththames.nihr.ac.uk/nihr-clahrc_north-thames-academy/self-management-child-youth-mental-health/. A recent pioneering effort to consider mentalizing stigma also took the 'intrapsychic filtering system' of gender-nonconforming individuals as their focus: Scandurra, C., Dolce, P., Vitelli, R., Esposito, G., Testa, R. J., Balsam, K. F., and Bochicchio, V. (2020). 'Mentalizing Stigma: Reflective Functioning as a Protective Factor against Depression and Anxiety in Transgender and Gender-Nonconforming People'. *Journal of Clinical Psychology*, 76(9): 1613–1630.

regarded social justice and mental health as distinct matters, making him hesitant to comment on issues of social justice as outside his expertise. However, he acknowledged that the #BlackLivesMatter campaign had brought him to re-evaluate this position: 'seeing humanity denied changes the way we all think.'⁶⁷ He subsequently issued a public statement: 'To pretend that the problems of disadvantage, of racism and of difference belongs to others is to allow them to continue unchallenged. Discrimination is totally antithetical to the core values of the Anna Freud Centre and we pledge our solidarity with our BAME colleagues, clients and students. As an organisation committed to ensuring that all infants, children and young people achieve their potential, we must address inequality based on ethnicity and offer support.'⁶⁸ It will be interesting to see what further ramifications may stem from Fonagy's reevaluation of the relevance of social criticism over the coming years.⁶⁹

To date, however, by far the primary intersection between the individual and society addressed by Fonagy and colleagues has been schools: 'we have focused on the question of how we can influence the environments that children inhabit. And of course, the crucial environment that comes to mind is the school setting.'⁷⁰ As we saw in Chapter 2, in Fonagy's biography, experiences of being bullied in secondary school was an important contributor to his depression, after he had been sent away by his parents to live in London; by contrast, experiences of validation and self-actualization in higher education helped reduce feelings of depression and suicidality. Another tributary to an interest in schools in buffering adversity likely stemmed also from the Anna Freudian tradition. Allison and Campbell observe that the concern of Fonagy and colleagues for education and mental health is 'in keeping with the tradition Anna Freud established of creating settings—such as the Matchbox School in Vienna in the 1920s and nursery schools or residential homes in the UK—whose ecology is conducive to growth and psychological development.'⁷¹ A further background influence has subsequently been identified by Fonagy, who has noted that his own attention to schools has occurred in the context of the growing sociological importance of schools in the late twentieth century, as children spend less and less time with their parents and more with their peers, so that 'children have become socialising agents for other children.'⁷² This is a role

⁶⁷ Fonagy, P. (2020). Tweet, 3 June. Accessed at: <https://twitter.com/peterfonagy/status/1268186685965242368?s=11>: 'It should not "just be victims of injustice and dehumanization who bear responsibility of continuing to mentalize"'

⁶⁸ Fonagy, P. (2020). 'Statement from Peter Fonagy on the Murder of George Floyd'. Accessed at: <https://www.annafreud.org/insights/news/2020/06/statement-from-peter-fonagy-on-the-murder-of-george-floyd/>.

⁶⁹ Already a few weeks later, in a virtual talk, Fonagy offered reflections on the danger of collective hypervigilance resulting from 'cultures of competition'—for instance, in the academic sphere in which individuals often receive rewards rather than teams, but also more widely in education and society. He explicitly advocated that such a shift in culture requires social and policy change, not solely adaptations in individual mentalizing. Fonagy, P. (2020). 'Trust and Interdependence: Lessons from the Study of Human Attachments', UCL Minds Event, virtual talk, 25 June.

⁷⁰ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre. See also Young-Bruehl, E. (2008). *Anna Freud: A Biography*, New Haven: Yale University Press. Emphasis on schools has been the most long-standing and significant area of focus beyond the individual for Fonagy and colleagues, with the exception of the family. However, in recent years, Fonagy has increasingly acknowledged the potential contribution of other institutions in prevention. Interview with Peter Fonagy cited in Maddox, L. (2018). *Blueprint: How Our Childhood Makes Us Who We Are*, London: Robinson, Chapter 7: 'I think mental health is inappropriately and inaccurately seen as just a subjective process, and I think making schools more mentalising, making workplaces more mentalising, making families more mentalising, making pubs more mentalising. I think this is where it's at.'

⁷¹ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

⁷² Fonagy, P. (2018). 'Peter Fonagy: Combating a Mental Health Crisis'. Accessed at: <https://www.goldmansachs.com/insights/talks-at-gs/peter-fonagy.html>: 'Since the second world war, since secondary education became universal, children are being educated in very large classrooms, in very large schools. This has had all kinds of advantages. But one of the things that it led to is that the number of hours of contact between adults and children

that Fonagy feels has been increasingly recognized during the COVID-19 pandemic during periods when children have not been able to access schools except remotely.⁷³

More proximally, for Fonagy the perception of schools as a critical site of preventative intervention appears to have stemmed from his collaboration with Stuart Twemlow in the early 2000s on the Peaceful Schools Project.⁷⁴ Twemlow, Fonagy, and colleagues began by thinking about secure, avoidant, and anxious/ambivalent school social systems.⁷⁵ These schools might look much the same when things are calm. But when an external or internal event presents a challenge to the school, this challenge is handled quite differently. In a secure school, the challenge is neither exaggerated nor minimized. There is collaboration in approaching the problem, making space for both individual feelings and evaluations of the situation on the way towards a well-coordinated collective response. In an avoidant school, the primary response to the challenge is denial and a sense of false bravado. The problem cannot be communicated about; instead, individuals in the school seem preoccupied with other tasks or difficulties they face. In an ambivalent/resistant school, challenges prompt a sense of ready panic, but without clear lines of communication for responding to the problem collectively or constructively. What hierarchy there is in the school tends not to be functional, and individuals may even undermine it further in their attempts to have their concerns heard. The school is also characterized by confusion between discipline and other domains such as relationships and safety, making it difficult to distinguish between more or less serious challenges, or how to judge a measured response to them. The result is a school in which there is a high level of expressed affect at the slightest provocation.

The extrapolation of the Ainsworth attachment classifications to the social system of the school by Twemlow, Fonagy, and colleagues entailed identification of common elements. Like the infant of a secure dyad in the Strange Situation, in the secure school there is room for expressions of distress, which are not transformed by minimizing or maximizing strategies, and the expectation of collaboration to resolve the problem. Like the infant of an insecure dyad in the Strange Situation, in the insecure school there is a transformation of information about the problem, so that it either may not be acknowledged as a source of concern or becomes a preoccupation to the exclusion of other kinds of priorities.⁷⁶ The innovative aspect of this picture was that a secure school was not conceptualized merely as the aggregate of secure or insecure individuals. Instead, the school was envisaged as having

has actually dropped. So children have become socialising agents for other children, and that's not how we were designed by evolution. We were designed by evolution to be looked after by people wiser than ourselves.'

⁷³ Fonagy, P. and Smith, J. (2020). 'Wellbeing: The Impact of the Pandemic and what Schools can do to Support Mental Health'. *Trust*, December 2020, <https://trust-journal.org.uk/current-edition/wellbeing-the-impact-of-the-pandemic-and-what-schools-can-do-to-support-mental-health/>

⁷⁴ This work can be seen in the context of a more general policy and research trend towards concern with the role of schools in mental health promotion. Durlak, J., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., and Schellinger, K. B. (2011). 'The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions'. *Child Development*, 82: 405–432; Weare, K. and Nind, M. (2011). 'Mental Health Promotion and Problem Prevention in Schools: What Does the Evidence Say?'. *Health Promotion International*, 26: 29–69.

⁷⁵ Twemlow, S. W., Fonagy, P., and Sacco, F. C. (2002). 'Feeling Safe in School'. *Smith College Studies in Social Work*, 72: 303–326.

⁷⁶ Due to a common basis in an information-processing model, there are significant analogies here to Hinshelwood's characterization of three kinds of group culture based on three kinds of relationship with the truth: the flexible, the rigid, and the fragmented. Hinshelwood, R. D. (1994). 'Attacks on the Reflective Space', in V. Shermer and M. Pines (eds), *Ring of Fire: Containing Primitive Emotional States*, London: Routledge, pp. 86–106. There is no indication that Twemlow and Fonagy knew of Hinshelwood's prior work.

dynamic properties at the level of the social system characteristic of security or insecurity.⁷⁷ Building on certain threads in the ideas of Main and Crittenden, security was elaborated as a state of non-distortion of information about threat and protection, whereas insecurity reflected certain transformations of this information (see Chapter 7). In general, Fonagy theorized that 'in secure organizations social influence is exercised in a framework dominated by an awareness of the mental states, concerns, thoughts, and feelings of individuals within the system'.⁷⁸ He also saw that in other organizations, too, the dynamics characteristic of insecurity could be seen. In fact, he remarked that 'many of us work in organizations where bullying is as, if not more, pervasive than at the average elementary school'.⁷⁹ An individual may have a secure or insecure attachment to individual caregivers, but Fonagy has argued that 'being responded to sensitively most of the time by many people fosters trust in the availability of the entire network', and that in such contexts, 'secure attachment is to a system rather than an individual'.⁸⁰ The idea of the externalization of the alien self was not invoked in conceptualizing these dynamics. This may reflect that the target audience for the work by Fonagy, Twemlow, and colleagues, likely educationalists and policy makers, may have been sceptical of this rather speculative and unobservable mechanism. However, it also perhaps reflects a decline in the frequency of appeal to the concept of the 'alien self' over the course of the 2000s.⁸¹

In 2009, Fonagy, Twemlow, and colleagues reported from a cluster randomized control trial of a mentalization-based intervention with school communities compared with pastoral-therapeutic interventions solely offered to disruptive students. The mentalization-based intervention had four major components. A first was use of reflective classroom discussions. For instance, 15 minutes were set aside at the end of each day for classes to reflect together on bully-victim-bystander relationships according to a structured format. Poster campaigns, stickers, and badges were used 'to create a climate where feelings were labelled and distress was acknowledged as legitimate'.⁸² A second component was a classroom management plan, which supported teachers in understanding and correcting problems at the root in the social system of the classroom, rather than punishing individual children who misbehave. Consideration of role of bystanders in the facilitation or inhibiting of bullying was especially integral to the programme. A third component was teaching pupils

⁷⁷ See also the discussion of social cohesion as a property of secure communities irreducible to the properties of the individuals that constitute it in Fonagy, P. and Higgitt, A. (2007). 'The Early Social and Emotional Determinants of Inequalities in Health', in G. Baruch, P. Fonagy, and D. Robins (eds), *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, Chichester, UK: John Wiley & Sons, pp. 3–34, p. 23.

⁷⁸ Fonagy, P. (2003). 'The Violence in our Schools: What can a Psychoanalytically Informed Approach Contribute?'. *Journal of Applied Psychoanalytic Studies*, 5(2): 223–238, p. 223.

⁷⁹ Ibid. 225. Fonagy has been at University College London straight through from his undergraduate degree, so he may be referring to the way that this university operates at times. However, remarks elsewhere lead to the suspicion that he is referring more to the Institute of Psychoanalysis, e.g. Fonagy, P. (2009). 'When analysts need to retire: The taboo of ageing in psychoanalysis', in B. Willock, R. Curtis, and L. Bohm (eds), *Taboo or not Taboo? Forbidden Thoughts, Forbidden Acts in Psychoanalysis and Psychotherapy*, London: Karnac Books, pp. 209–227.

⁸⁰ Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>; Slides. Accessed at: <https://www.borderlinepersonalitydisorder.org/wp-content/uploads/2019/02/Podcast-2.24.19-compressed.pdf>.

⁸¹ The concept of projective identification is used by Twemlow when writing for a psychoanalytic audience at the start of the decade. Twemlow, S. W. (2000). 'The Roots of Violence: Converging Psychoanalytic Explanatory Models for Power Struggles and Violence in Schools'. *Psychoanalytic Quarterly*, 69(4): 741–785.

⁸² Fonagy, P., Twemlow, S. W., Vernberg, E. M., Nelson, J. M., Dill, E. J., Little, T. D., and Sargent, J. A. (2009). 'A Cluster Randomized Controlled Trial of Child-Focused Psychiatric Consultation and a School Systems-Focused Intervention to Reduce Aggression'. *Journal of Child Psychology and Psychiatry*, 50(5): 607–616, supplemental materials.

nonaggressive physical and cognitive strategies for when they find themselves a bystander or a victim of bullying. Role play was used to help pupils see that these were structural roles, and could in principle be populated by anyone, but that the roles could be handled in various ways. A fourth component was the introduction of a peer mentoring system.

Both the mentalization-based intervention and the intervention focused on disruptive students were greeted enthusiastically by teachers. However, only the mentalization-based intervention proved effective. There were reductions in peer-reported aggression, peer-reported victimization, support by bystanders for aggressive behaviour, and an increase in concern about the experiences of victims of bullying. In the classroom, the schools that implemented the mentalization-based approach saw medium to large reductions in pupils off-task and disruptive behaviour by children as measured by researchers' observations. Fonagy, Twemlow, and colleagues observed, from a review of the literature, that these behavioural changes were the largest seen to date in any randomized trial of a school-based intervention.⁸³

The collaboration with Twemlow and related thinking about schools as social systems led Fonagy and colleagues to quite a distinctive position on the role of schools in society. This was reflected in the influential THRIVE model developed by Wolpert, Fonagy, and colleagues for the provision of child and adolescent mental health.⁸⁴ THRIVE is the operating model used for 47% of the 0–18 population in England, and is recommended in the NHS Long Term Plan.⁸⁵ Wolpert, Fonagy, and colleagues divided provision into distinct categories, with characteristic providers, goals, and discourses. Their goal was to structure services around the different strategies for providing help, rather than severity of need or diagnosis. This aimed to help facilitate shared decision making with young people about what strategies they think will be most relevant and useful. A first level of support was to help young people cope and to foster their resilience to challenges. It was argued that 'wherever possible, this provision should be provided within education or community settings, with education often (though not always) the lead provider and educational language (a language of wellness) as the key language used.'⁸⁶ A further level of support offered children help with mental health difficulties: 'provision for this group should be provided with health as the lead provider and using a health language (a language of treatment and health outcomes)'. A distinctive aspect of the THRIVE approach to help was that 'treatment would involve explicit agreement at the outset as to what a successful outcome would look like, how likely this was to occur by a specific date, and what would happen if this was not achieved in a reasonable timeframe.'⁸⁷

Support for coping and help with mental health difficulties were contrasted by Wolpert, Fonagy, and colleagues with risk support. Here 'social care may often be the lead agency and

⁸³ Ibid.: 'Medium to strong effects for classroom observations (mean ES = .97, range: .84–1.1). Past research on school-wide multimodal interventions has only demonstrated modest effects on aggression against TAU [treatment as usual] (Wilson et al., 2003). Medium effect sizes are normally only associated with intense studies with small samples (Wilson et al., 2001; Smith et al., 2003). CAPSLE produced a number of modest and some large effect sizes, particularly with behavioral measures. The findings are also notable because the sample was not particularly high-risk and high-risk samples normally generate larger effects. (Mytton et al., 2002; Wilson et al., 2001, 2003)' (p. 614).

⁸⁴ Wolpert, M., Harris, R., Jones, M., ... Fonagy, P. (2014). 'THRIVE The AFC–Tavistock Model for CAMHS'. Accessed at: http://www.annafreud.org/media/2552/thrive-booklet_march-15.pdf.

⁸⁵ National Health Service (2019). *The NHS Long Term Plan*. Accessed at: <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf>.

⁸⁶ Wolpert, M., Harris, R., Jones, M., ... Fonagy, P. (2014). 'THRIVE The AFC–Tavistock Model for CAMHS'. Accessed at: http://www.annafreud.org/media/2552/thrive-booklet_march-15.pdf.

⁸⁷ Ibid.

the language of social care (risk and support) is likely to be dominant'. The authors acknowledged that 'perhaps the most contentious aspect of the THRIVE model' is that services should acknowledge that not every child can be helped through mental health provision, and that only risk management rather than mental health treatment should be provided if there is reason to believe that mental health treatment would be ineffective. The goal of support for coping should be to improve young people's resilience; the goal of help should be to reduce specific identified symptoms; and the goal of risk support should be to reduce young people's risk of harm to themselves or others. In the THRIVE model, then, schools are given a critical role. When more complex needs emerge, specialist mental health services and/or social care might become involved. But schools are depicted by THRIVE as holding fundamental responsibility as the locus of primary support for young people by social systems beyond their families. Rather than prioritizing the expansion of specialist mental health provision, the THRIVE model suggests taking the school community as the fundamental target for intervention.

Fonagy and colleagues provide the example of the approach they would ideally like to see for 'Billy', an 11-year-old frequently found by his teacher 'Ms. Jones' to be disruptive in class, and to engage in power-struggles and bullying towards other students:

Would Billy benefit from individual therapy? Experience shows that boys like Billy respond poorly to such efforts, however skilled, if they do not take place in the context of concurrent family and social interventions. We feel that disrupting the vicious cycle that a child such as Billy finds himself in should be undertaken in school. Furthermore, it may be best, given Billy's sensitivity to humiliation, if the intervention does not directly concern Billy at all but, rather, the whole class ... The procedure in the context of our program would have been for Ms. Jones to stop the class immediately after Billy started creating a commotion, and mark some space for reflection on what was happening. In other words, she would recognise the assault on her own mentalising that she was experiencing ... in the process seek to draw upon the collective mentalising capacity of her classroom to create an environment in which Billy's arousal could reduce.⁸⁸

There appears to have been a reciprocal support between Fonagy's interest in schools and his growing attention to epistemic trust and learning. On the one hand, the idea of epistemic trust as an openness to learning from others is well encapsulated by the school context.⁸⁹ On the other hand, the increased focus on epistemic trust has further solidified the conviction of Fonagy and colleagues that schools are a privileged locus for preventative work in public mental health, especially where teachers can mentalize and offer ostensive cues to pupils.⁹⁰

⁸⁸ This passage appears in Twemlow, S. W., Fonagy, P., and Sacco, F. C. (2005). 'A Developmental Approach to Mentalizing Communities: I. The Peaceful Schools Experiment', *Bulletin of the Menninger Clinic*, 69(4): 282–304, p. 289; and is repeated in Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 447–448. It might be argued that this strategy places Billy's mentalizing and learning above the curricular learning of the class as a whole, whose lesson has to be paused to reflect on Billy's feelings. It can be anticipated that Fonagy and colleagues would counter that Billy's behaviour is already disrupting class learning, and that the opportunity for the class to use and develop collective mentalizing capacity will benefit their well-being, and—possibly—their future curricular learning. It is to be hoped that mentalization-based interventions in schools may be able to explore such questions.

⁸⁹ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication', *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9. See also Lane, J. D. and Harris, P. L. (2015). 'The Roles of Intuition and Informants' Expertise in Children's Epistemic Trust', *Child Development*, 86(3): 919–926.

⁹⁰ Fonagy, P. (2019). 'Attachment-Aware Teaching with Peter Fonagy', *Podogogy*, 7(5): <https://play.acast.com/s/tes-the-education-podcast/32054866-4898-47ce-b70d-3995276773a9>; Fonagy, P. (2016). 'Keynote Address.'

Furthermore, the growing momentum at the Anna Freud Centre around using schools as a site for scalable public mental health intervention has led to major ongoing trials of various interventions, led by the UCL-Anna Freud Evidence-Based Practice Unit.⁹¹ This includes the 'Education for Wellbeing' trial, which will recruit 26,000 pupils across England to assess the relative effectiveness of five different school-based interventions.⁹² The Anna Freud Centre has also pursued research on strategies to support the mental health well-being of school staff, as part of facilitating mental health within schools.⁹³

In recent years especially, Fonagy has been proactive in advocating to policy makers regarding individual self-regulation and schools as potential sites for public mental health preventative work.⁹⁴ In relation to individual self-regulation, Fonagy has been critical of the effectiveness of specialist mental health services, both in terms of reach and in terms of aggregate clinical impact.⁹⁵ So, instead of advocating for the resourcing of specialist mental health services, he has been a national policy leader in the expansion of primary care mental health services for children and adolescents, within the 'Improving Access to Psychological Therapies' framework. This included various provisions that could be delivered within a primary care setting, such as brief psychological therapies for anxiety and depression for children, and parenting classes when young children have conduct problems.

Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde, 24 November, Berlin. Accessed at: <https://twitter.com/TheLancetPsych/status/801813352515272705>

⁹¹ E.g. Wolpert, M., Humphrey, N., Deighton, J., Patalay, P., Fugard, A. J., Fonagy, P., ... and Panos, V. (2015). 'An Evaluation of the Implementation and Impact of England's Mandated School-Based Mental Health Initiative in Elementary Schools'. *School Psychology Review*, 44(1): 117–138.

⁹² The five school-based interventions are:

1. 'A set of five lessons for Y9 that use role play designed to improve pupils' understanding of mental health and reduce suicide rates. Developed in Sweden and America, Youth Aware Mental Health (YAM) encourages pupils to share their own ideas about how to maintain good mental health and how to help each other to find ways to resolve everyday dilemmas
2. A teacher training programme developed in Canada called *The Guide*. Adapted for England for the study, it develops teachers' understanding of mental health, trains them on how to teach their pupils about it and addresses stigma
3. A series of eight lessons designed to increase young people's skills around personal safety and managing their mental health, as well as helping them to identify their support networks
4. Training pupils in relaxation techniques embedded into the school day, every day for five minutes
5. Training pupils in mindfulness embedded into the school day, every day for five minutes.'

Accessed at: <https://www.ucl.ac.uk/evidence-based-practice-unit/research/research-projects>.

⁹³ Garland, L., Linehan, T., Merrett, N., Smith, J. and Payne, C. (2018). *Ten Steps Towards School Staff Wellbeing*. London: Anna Freud Centre. Accessed at: <https://www.annafreud.org/media/8506/school-staff-wellbeing-report-final-corrected-512.pdf>.

⁹⁴ Fonagy, P., and Higgitt, A. (2004). 'Early Mental Health Intervention and Prevention: The Implications for Government and the Wider Community', in B. Sklarew, S. W. Twemlow, and S. M. Wilkinson (eds), *Analysts in the Trenches: Streets, Schools, War Zones*, Mahwah, NJ: Analytic Press, pp. 257–309.: 'Psychoanalysts have shrunk from involvement with public policy. The nature of a psychoanalyst's commitment to an individual, the intensely personal nature of the interaction, which sustains the "no holds barred" study of subjectivity, is experienced by most analysts as incompatible with commitment to populations, institutions, organisations and other faceless systems. This attitude, of course, is illogical' (p. 257). On schools as a locus for intervention, see Gee, B., Wilson, J., Clarke, T., Farthing, S., Carroll, B., Jackson, C., ... and Nottley, C. (2020). 'Delivering Mental Health Support within Schools and Colleges—A Thematic Synthesis of Barriers and Facilitators to Implementation of Indicated Psychological Interventions for Adolescents'. *Child and Adolescent Mental Health*, 26(1): 34–46.

⁹⁵ Fonagy, P. and Clark, D. M. (2015). 'Update on the Improving Access to Psychological Therapies Programme in England: Commentary on ... Children and Young People's Improving Access to Psychological Therapies'. *BJPsych Bulletin*, 39(5): 248–251, citing Warren, J., Nelson, P., Mondragon, S., Baldwin, S., and Burlingame, G. (2010). 'Youth Psychotherapy Change Trajectories and Outcomes in Usual Care'. *Journal of Consulting and Clinical Psychology*, 78(2): 144–155; Fonagy, P., Pugh, K., and O'Herlihy, A. (2017). 'The Children and Young People's Improving Access to Psychological Therapies (CYP IAPT) Programme in England', in D. Skuse, H. Bruce, and L. Dowdney (eds), *Child Psychology and Psychiatry: Frameworks for Clinical Training and Practice* (3rd edn), Oxford: John Wiley & Sons, pp. 429–435.

The exact nature of Fonagy's policy work is difficult to document, because there is not much information in the public domain. In the context of major cuts to mental health services since 2011, as part of the government's austerity agenda, Fonagy reports his sense that his policy engagement has contributed towards the direction of public resources towards child and adolescent mental health:

I think it's the thing I've done that I'm most proud of. The service was not fit for purpose. In many parts it may still not be. In a situation where [mental health services for children and adolescents] suffered a financial penalty greater than at any time in its history, it has not collapsed ... in the last budget of this government, £1.25 billion was set aside for children's mental health—£250 million a year over the next five years. That's getting back to the level it was at before the cuts began.⁹⁶

Under the coalition government, Fonagy served as an academic adviser to Paul Burstow and Norman Lamb, Ministers of State for Care.⁹⁷ In 2013, his work at the research-policy interface led to the award of an OBE by the Queen. And he was singled out by the Children's Minister in 2014 for praise for his achievements in policy-relevant research.⁹⁸

In 2017, Fonagy and colleagues were asked by the Department of Health to conduct a 'systematic review of the evidence relating to the mental health of children and young people'. The systematic review and its conclusions have not been published, though future publication has been promised.⁹⁹ However, the Department of Health Green Paper stated that Fonagy and colleagues had advocated for greater use of schools for preventing, identifying, and responding to mental health issues:

The school environment is well suited to a graduated approach to children's mental health, where children at risk can be identified and interventions can be offered to address problems. As the school environment can present triggers for many difficulties (such as social anxiety), it is therefore also a good place to find support to manage them.¹⁰⁰

The central innovation in the Green Paper was the creation of new Mental Health Support Teams attached to schools, with around 8,000 staff. This would be larger than the whole of the existing children and young people's mental health services workforce in the NHS. It was

⁹⁶ Fonagy, P. (2015). 'Peter Fonagy on Psychoanalysis and IAPT' The History of Emotions Blog, posted on 14 May by Jules Evans. Accessed at: <https://emotionsblog.history.qmul.ac.uk/2015/05/peter-fonagy-on-psychoanalysis-and-iapt/>. See also Pickersgill, M. (2019). 'Access, Accountability, and the Proliferation of Psychological Therapy: On the Introduction of the IAPT Initiative and the Transformation of Mental Healthcare'. *Social Studies of Science*, 49(4): 627–650. More recently, Fonagy has also collaborated on an open letter to the government, advocating for greater allocation of funding for children's well-being and mental health in the context of COVID-19. Bajawa, S., Thomas, E., Fonagy, P. et al. (2020). 'Open Letter to Government on Young People's Mental Health During Lockdown'. Accessed at: <https://www.bps.org.uk/news-and-policy/open-letter-government-young-peoples-mental-health-during-lockdown>.

⁹⁷ University College London (2014) REF 2014, Psychology, Psychiatry, and Neuroscience: Environment Statement. Accessed at: [https://results.ref.ac.uk/\(S\(jj2mvvb3fbee3zpqb1artx2d\)\)/DownloadFile/EnvironmentTemplate/PDF?subId=1075](https://results.ref.ac.uk/(S(jj2mvvb3fbee3zpqb1artx2d))/DownloadFile/EnvironmentTemplate/PDF?subId=1075).

⁹⁸ <https://www.gov.uk/government/speeches/edward-timpson-on-better-support-for-vulnerable-adolescents>.

⁹⁹ Secretary of State for Health and Social Care and the Secretary of State for Education (2018). 'Government Response to the First Joint Report of the Education and Health and Social Care Committees of Session 2017–19 on *Transforming Children and Young People's Mental Health Provision: A Green Paper*', p. 10. Accessed at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/728902/HESC_Print_3_.pdf.

¹⁰⁰ Department of Health (2017). *Transforming Children and Young People's Mental Health Provision: A Green Paper*, London: HMSO, p. 10.

also expected that schools would liaise more regularly with mental health services. The Anna Freud Centre would contribute to the training of this workforce through a new Diploma for Mental Health Practitioners in Schools. Fonagy has worked with the Secretary of State for Health, Jeremy Hunt, to argue publicly for the benefits of a school-focused approach to improving young people's mental health.¹⁰¹ The Education and Health Select Committees have criticized the Green Paper, alleging that it passes responsibility for children's mental health to schools, rather than offering either adequate resources to individual mental health provision or social change to address the factors at the root of young people's suffering.¹⁰² The Committees also expressed concern that young people with potential additional mental health needs, such as those in care or on the edge of care, are given insufficient priority in the reforms. No doubt Fonagy would have wished that more resources were available. However, for decades he has consistently assumed public resources to be limited, and on this basis argued for a prioritization of primary prevention in schools, and a comparatively reduced prioritization of secondary mental health care. He has also been a consistent advocate for greater interagency working between sectors.¹⁰³

Collective mentalizing capacity

In 2017, Di Stefano and colleagues alleged that it 'seems that Fonagy and colleagues have paid little attention to the context in which the relational, mental, and affective experiences of the individuals originate and consolidate'. In doing so, Fonagy and colleagues fail to consider the 'rules, meanings, and shared models of behavior that produce relationships and social exchanges capable of supporting or, on the contrary, hindering processes of exploration and meaning that individuals attribute to their own (and others) actions'.¹⁰⁴ As we saw in Chapter 3, when Fonagy was initially elaborating the concept of mentalization in the early 1990s on the basis of existing work in social cognition, he included some elements of social cognition and not others. An important exclusion was that cognition about social norms was not taken up as part of the concept of mentalizing. In contrast to social cognition, then, mentalization was not concerned with norms about inequality, hierarchy, and intrinsic differences between groups that may play a significant role in oppression. Fonagy's assumption

¹⁰¹ Department of Health and Social Care (2018). 'Children and Young People's Mental Health Green Paper': webcast with Jeremy Hunt. Accessed at: <https://www.youtube.com/watch?v=j4gAI402ndM>. Fonagy is also co-signatory on a public letter advocating that assessment of schools should include appraisal of their provision for mental health. See Turner, C. (2018). 'Ofsted is Considering New Assessment to Ensure Schools Look after Pupils' Mental Health and Wellbeing', 22 December. Accessed at: <https://www.telegraph.co.uk/education/2018/12/22/ofsted-considering-new-assessment-ensure-schools-look-pupils/>.

¹⁰² Education and Health Select Committees (2018). *The Government's Green Paper on Mental Health: Failing a Generation*, London: HMSO.

¹⁰³ The Anna Freud Centre has been funded to evaluate aspects of the new approach. This work is largely still ongoing, but in an early pilot of the new approach mental health leads from 255 schools took part in two-day workshops focused on interagency working. A study conducted by Cortina and colleagues at the Anna Freud Centre found that 55% of school mental health leads reported being in 'monthly' or 'continuous' contact with mental health services at follow-up, compared with a quarter at baseline. Limitations of the study should be noted: the data comes from self-report, and less than a third of participants completed the survey at both time points. Cortina, M. A., Shipman, J., Saunders, F., Day, L., Blades, R., Smith, J., and Wolpert, M. (2019). 'Embedding Interagency Working between Schools and Mental Health Specialists: A Service Evaluation of the Mental Health Services and Schools and Colleges Link Programme Workshops'. *Clinical Child Psychology and Psychiatry*, 24(4): 906–920.

¹⁰⁴ Di Stefano, G., Piacentino, B., and Ruvalo, G. (2017). 'Mentalizing in Organizations: A Psychodynamic Model for an Understanding of Well-Being and Suffering in the Work Contexts'. *World Futures*, 73(4–5): 216–223, p. 218–219. See also Sperry, M. (2013). 'Putting our Heads Together: Mentalizing Systems'. *Psychoanalytic Dialogues*, 23(6): 683–699.

in the early 1990s appears to have been that the key concerns of mentalisation—intentions, beliefs, affects—can only belong to individual subjects, and cannot belong to groups.

Later in the 1990s, when working on the reflective functioning scale with Howard and Miriam Steele and Anna Higgitt, Fonagy treated sociological discussions that place attachment figures in their cultural and historical context not simply as irrelevant to mentalizing, but specifically as bad mentalizing: ‘Passages rated as “1” may be sociological, excessively generalised, concrete or overwhelmingly egocentric.’¹⁰⁵ They also advised coders to give low scores to accounts of mental states that seemed drawn from ‘shared culture’. It was assumed that ‘shared culture’ was opposed to the individual’s own sense of mental states as ‘personally significant and meaningful’.¹⁰⁶ Across the reflective functioning scale manual, what references to social and cultural factors occur seem premised on the expectation of a trade-off between psychological-mindedness and sociological- or group-mindedness.¹⁰⁷ This was a tendency supported by the poorly articulated distinction between ‘not mentalizing’ and ‘non-mentalizing’, sometimes leading one to be mistaken for the other. It may also have been supported by Fonagy’s distrust of the dynamics of large groups, which he regarded as tending away from acknowledgement of individual experiences.¹⁰⁸ In the 2000s, Fonagy allowed that it would be possible to mentalize non-corporeal individuals, such as ‘explicating the mental states of story characters’ or the intentions of divine beings.¹⁰⁹ However, groups were beyond the pale.¹¹⁰

¹⁰⁵ Fonagy, P., Target, M., Steele, H., and Steele, M. (1998). *Reflective Functioning Manual, Version 5*, London: UCL/Anna Freud Centre, p. 28.

¹⁰⁶ Ibid.: ‘The passage may be “marked” in reflective-functioning because the view of mental states presented by the subject is unusual and surprising to the rater. Passages which cast an original perspective, which nevertheless is readily understandable to the rater, reflect mentalization on the part of the subject. Raters should however be aware of the possibility of “borrowed” reflective-functioning, where the subject is repeating ideas presented to him/her in other contexts. In such instances a rating of “3” would probably be more appropriate’ (p. 29); ‘the passage must contain sufficient “surprise” and coherence for the rater to feel it is unlikely to have come from contaminating sources. The passage should therefore have a personal character, i.e., experienced as personally significant and meaningful, and may seem to be developing further during the interview itself’ (p. 30); ‘In contrast to interviews rated “0”–“4”, interviews rated “5” give convincing indications to the rater that speakers have some kind of a model of the mind of attachment figures as well as a model of their own mind which is relatively coherent even if it is simple, and is unlikely to have been solely derived from shared culture rather than from personal experience’ (p. 34).

¹⁰⁷ To take another example of this assumption: Luyten, P., Fonagy, P., Lowyck, B., and Vermote, R. (2012). ‘The Assessment of Mentalization’, in A. W. Bateman and P. Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice*, Washington, DC: American Psychiatric Publishing, pp. 43–66: ‘What does bad mentalisation look like? ... Focus on external factors (e.g. government, school, colleagues, neighbors)’ (p. 59).

¹⁰⁸ E.g. Fonagy, P. (2003). ‘The Violence in our Schools: What can a Psychoanalytically Informed Approach Contribute?’, *Journal of Applied Psychoanalytic Studies*, 5(2): 223–238: ‘Coercion is the rule within most social systems.’ Ratner, R. H. (2014). ‘Interview with Peter Fonagy’, in *Borderline*. Master of Fine Arts, The University of Texas at Austin. Accessed at: <https://repositories.lib.utexas.edu/bitstream/handle/2152/28667/RATNER-MASTERSREPORT-2014.pdf?sequence=1>. From interview with Fonagy: ‘Mentalization is one of best established of human capacities, can break in all of us. It is fairly predictable when it can break: large groups (hard to mentalize when we’re in a large crowd, as a consequence, large crowds are dangerous, you can forget what someone is feeling hence can perform atrocities’ (p. 123).

¹⁰⁹ E.g. Fonagy, P., Target, M., Gergely, G., Allen, J. G., and Bateman, A. W. (2003). ‘The Developmental Roots of Borderline Personality Disorder in Early Attachment Relationships: A Theory and Some Evidence.’ *Psychoanalytic Inquiry*, 23(3): 412–459, p. 433; Fonagy, P. (2009). ‘Commentary on “Forgiveness”’, in S. Akhtar (ed.), *Good Feelings: Psychoanalytic Reflections on Positive Emotions and Attitudes*, London: Karnac/International Psychoanalytic Association, pp. 411–452, p. 423. Fonagy would later contribute to work in which the capacity to consider the mental states of fictional individuals was specifically used to operationalize mentalizing. See Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., and Somma, A. (2018). ‘Thinking about Assessment: Further Evidence of the Validity of the Movie for the Assessment of Social Cognition as a Measure of Mentalistic Abilities.’ *Psychoanalytic Psychology*, 35(1): 127–141.

¹¹⁰ This may have been reinforced by Gergely’s account of marked mirroring and ostensive cues as conveying acknowledgement of individual agency, a property that may have been assumed to only characterize individuals.

Fonagy and Target have subsequently regretted their early rejection of sociological and group-focused thinking. In fact, Target has substantially redirected her approach over the past 15 years towards greater engagement with sociological theory and methods.¹¹¹ Fonagy and Target have recognized that 'every patient comes to an assessment with his own language and frame of reference for emotional distress, with his own theories consonant with cultural idioms for the expression of emotional distress'.¹¹² In the landmark 'What We have Changed our Minds About' paper from 2017, Fonagy and colleagues stated that they had changed their minds 'about the relationship between the individual and culture', seeing the need for a 'more systemic, less intrapsychic approach'.¹¹³

As we saw above, in the 2000s, Fonagy, Twemlow, and colleagues struggled to explicitly work out a basis for extrapolating mentalization from interpersonal processes to social systems, as they had with security and insecurity. Mostly, they seem to have assumed that a capacity to conceive of and reconsider thoughts and feelings could only be a property of individuals, and could only take as its target other individuals: 'Culture is the product of individual minds. It's an aggregate. It is whatever comes through from what lots of minds have in common'.¹¹⁴ In such statements, no autonomy was ascribed to social systems as more than the aggregate of individuals. However, as we saw social environment, there were other times when Fonagy, especially in work with Twemlow, moved towards an account in which the properties of a community itself could i) be considered to facilitate mentalization, or ii) be taken as the target of mentalization by individuals concerned with perspectives and affects circulating in a social system.¹¹⁵ Both were reflected in the first two components of their Peaceful Schools intervention. The class as a whole were asked to reflect together on bully-victim-bystander relationships, supported by the teacher's understanding of and attention to these systemic dynamics. Teachers were encouraged to 'draw upon the collective mentalising capacity of [the] classroom'.¹¹⁶

What could the collective mentalizing capacity of the classroom be? The nature of the Peaceful Schools intervention implicitly suggests some aspects. One is dedicated time. The collective expectation of time to reflect at the end of the day might be part of the collective mentalizing capacity, alongside the accumulation of previous successful discussions as a

¹¹¹ Target, M., and Fonagy, P. (2003). 'Attachment Theory and Long-Term Psychoanalytic Outcome: Are Insecure Attachment Narratives Less Accurate?', in M. Leuzinger-Bohleber, A. U. Dreher, and J. Canestri (eds), *Pluralism and Unity? Methods of Research in Psychoanalysis*, London: International Psychoanalytical Association, pp. 149–167: 'We are gradually trying to feel our way towards an application of qualitative methodology from sociology' (p. 165).

¹¹² Lemma, A., Target, M., and Fonagy, P. (2011). *Brief Dynamic Interpersonal Therapy: A Clinician's Guide*, Oxford: Oxford University Press, p. 57.

¹¹³ Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9.

¹¹⁴ See also Fonagy, P. (1999). 'Interview with Peter Fonagy', in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98, p. 83.

¹¹⁵ Though not a discussion of the community as a target for mentalizing, a reflection on the community as a potential target for attachment (i.e. secure base/safe haven dynamics) was offered in Fonagy, P., Target, M., Steele, M., Steele, H., Leigh, T., Levinson, A., et al. (1997). 'Morality, Disruptive Behavior, Borderline Personality Disorder, Crime, and their Relationships to Security of Attachment', in L. Atkinson and K. J. Zucker (eds), *Attachment and Psychopathology*, New York: Guilford Press, pp. 223–274, p. 224: 'Unstable attachment histories [are] more common in these [dangerous] neighbourhoods, thus reducing the overall likelihood of strong attachments to the community' (p. 231).

¹¹⁶ This passage appears in Twemlow, S. W., Fonagy, P., and Sacco, F. C. (2005). 'A Developmental Approach to Mentalizing Communities: II. The Peaceful Schools Experiment'. *Bulletin of the Menninger Clinic*, 69(4): 282–304, p. 289; and is repeated in Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 447–448.

collective reference point.¹¹⁷ Another part of the scaffolding of collective mentalizing capacity may be space. That the daily discussions took place in a familiar classroom may suggest the physical space as a potential contributory to collective mentalizing capacity, as a secure base.¹¹⁸ The collective mentalizing capacity may encompass the circulation of images that make acknowledgement of feelings a taken-for-granted part of the institutional climate of the wider school. The school's coordination and coherence of approach to discipline might also be part of its collective mentalizing capacity. No doubt the safety of the school environment would also be a relevant collective resource for facilitating mentalizing. None of these features seem readily reducible to merely the aggregate of the mentalizing capacities of the school's members.

Nonetheless, the concept of a 'collective mentalizing capacity' remained underdeveloped, with three subterranean and partial exceptions. One partial exception has been thinking in the development and running of the Pears Family School within the Anna Freud Centre, founded in 2014. The school takes pupils excluded from or unable to cope with mainstream education. A requirement is that parents are willing to join a parent group and participate in the life of the school. The founding vision was for a community with 'many of the features of a mainstream school alongside the non-institutionalized values of a family environment'.¹¹⁹ It is the ambition that each child returns to mainstream education after receiving sufficient benefit from the Family School, and to date this has been the case for the majority of children.¹²⁰ Descriptions of the school do not make explicit mention of the collective mentalizing capacity institutionalized in the Pears Family School. However, at times, there seems implicit acknowledgement that mentalizing and epistemic trust in the school is more than the sum of these qualities in its members. One example is Fonagy's response when the Pears Family School and Anna Freud Centre won the Grand Prix Design Award in 2019: 'The space we have is one for bringing people together, and the environment encourages collaboration, creativity, and inclusivity'.¹²¹ This suggests acknowledgement of the role of physical space as a resource towards collective mentalizing capacity. Another example of implicit acknowledgement of collective mentalizing capacity is in a remark by Allison and Campbell about the ambitions of the school, that 'the work of building a community at the school and ensuring that it functions as a mentalizing environment is a critical part of its therapeutic

¹¹⁷ On how the benefits of group-level processes for reconsideration of thoughts may be frequently underestimated, see Mercier, H., Trouche, E., Yama, H., Heintz, C., and Giroto, V. (2015). 'Experts and Laymen Grossly Underestimate the Benefits of Argumentation for Reasoning'. *Thinking & Reasoning*, 21(3): 341–355. Mercier has also identified that the experience of a group majority is not reducible to the sum of the aggregates in contexts of social influence. Mercier, H. and Morin, O. (2019). 'Majority Rules: How Good are We at Aggregating Convergent Opinions?'. *Evolutionary Human Sciences*, 1. Mercier was one of Sperber's central collaborators in introducing the concept of epistemic vigilance (see Chapter 7).

¹¹⁸ The importance of the physical setting and a physical calendar as secure bases to facilitate young people's mentalizing has been emphasized by Midgley, N., Ensink, K. and Lindqvist, K. (2017). *Mentalization-based Treatment for Children: A Time-limited Approach*, New York: American Psychological Association.

¹¹⁹ The Family School (Anna Freud) (2014). Application for Free School status. Accessed at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/439138/Wave_4_-_The_Family_School.pdf.

¹²⁰ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre: '64 per cent of pupils who have attended the school since it opened have been successfully reintegrated into mainstream school within the school's own target of four terms. Of the pupils who have returned to mainstream schools since the school opened, 95 per cent have succeeded in retaining their school places.'

¹²¹ Accessed at: <https://www.annafreud.org/insights/news/2019/11/double-award-win-for-our-new-home-the-kantor-centre-of-excellence/>.

work ... The community of Pears Family School seeks to operate as a mentalizing training ground for children and families.¹²²

Another partial exception to the underdevelopment of the concept of collective mentalizing capacity has been the work of Asen and Fonagy, considering the applicability of mentalization-based family therapy. In two papers from 2017 on family violence, they argued that ‘the family consulting room can be a unique environment where the collective mentalizing capacity of the family, with the support of the therapist’s own reflective capacities, can be mobilized to identify the thoughts and feelings that might have triggered a problematic interaction.’¹²³ Implicit in their account is a distinction between individual mentalizing and collective mentalizing capacity. The former is defined as an ‘imaginative mental activity that entails perceiving and interpreting human behavior in terms of intentional mental states.’¹²⁴ The latter is left undefined, but appears from the way the term is used to be a property of the family system that permits the identification and reflection on thoughts and feelings of and between its members. The individual mentalizing skills of family members support the collective mentalizing capacity. But even when these skills fail, the collective mentalizing capacity may have an independent contribution to offer.

By the same token, individual mentalizing skills may have reduced effect, or could even backfire, depending on the state of the family’s collective mentalizing capacity. Asen and Fonagy did not explicitly theorize collective non-mentalizing processes operating at the level of the family system, and distinct from the non-mentalizing of individuals in the family. However, they took steps towards such an account. They stated that ‘a system—be that a family or other social group—that is characterized by blindness to the mental states of self and others will tend to create systems of social influence where coercion and humiliation play a key role.’¹²⁵ They discussed ‘escalating nonmentalizing interchanges’, in which the interactions of individuals in the family seem to become more than the sum of their parts in facilitating pretend mode, psychic equivalence, or teleological mode.¹²⁶ They also describe the way that violence within the family system can be stabilized or destabilized by use of externalization of the alien self by a member or members of the family. Yet, at the level of clinical technique, their approach is quite firmly oriented by a focus on helping family members understand their own minds and those of others. There is little attention at the level of technique, beyond encouraging individual mentalizing, to facilitation of collective mentalizing capacity, or reduction of collective non-mentalizing processes. Despite clear apparent relevance, it remains unexamined—for instance, whether the group dynamics that would facilitate automatic mentalizing would help or hinder controlled mentalizing, and vice versa.

A third partial exception to the neglect of collective mentalizing capacity is in work on group-based MBT, building on the long heritage of attention to group-level processes in group psychotherapy.¹²⁷ Occasionally, Fonagy and colleagues describe the need to

¹²² Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

¹²³ Asen, E. and Fonagy, P. (2017). ‘Mentalizing Family Violence Part 2: Techniques and Interventions.’ *Family Process*, 56(1):22–44, p. 29.

¹²⁴ Ibid. 6–21, p. 8.

¹²⁵ Ibid. 15.

¹²⁶ Ibid. 13. We have not found any point in Fonagy’s writings where he discusses Freud’s Dora case study. However, this seems exemplary in showing pretend mode, psychic equivalence, and teleological mode helping to sustain one another within the family system. For a related interpretation of Dora, see Billig, M. (1999). *Freudian Repression: Conversation Creating the Unconscious*, Cambridge: Cambridge University Press.

¹²⁷ Winship, G. (2003). ‘The Democratic Origins of the Term “Group Analysis”’: Karl Mannheim’s Third Way for Psychoanalysis and Social Science. *Group Analysis*, 36(1): 37–51; Roseneil, S. (2019). ‘Broader (than Psychoanalysis) and Deeper (than Sociology): The Psychosocial Promise of Group Analysis.’ *Psychoanalysis, Culture & Society*, 24: 493–501.

foster particular kinds of 'group culture' or 'a secure atmosphere within the treatment milieu'.¹²⁸ However, such claims are only ever made in passing. More generally, the collective mentalizing capacity of the MBT group has been treated as the sum of individual capabilities.¹²⁹ It is revealing that in Sigmund Karterud's manual for group-based MBT he specifically stated that the formation of group culture in contributing to or hindering mentalizing is beyond the scope of the manual.¹³⁰ At points, Karterud referred to Foulkes's concept of 'matrix', the communication network of the group materialized by the history of their interaction.¹³¹ He emphasized the importance of such collective processes in shaping the potential for mentalizing. However, he also urged that a distinction between MBT and many other forms of group therapy is that in the former the therapist should generally avoid discussion of group-level processes, because what matters is the individual's capacity to consider other individual minds and their own. Karterud also perceived a risk that discussion of group-level processes had special risks for entry into pretend mode. The MBT therapist may certainly think about group processes, but the primary way in which they engage with them is through 'stimulating and assisting the group in discussion of group relevant themes'.¹³² There is barely mention even of group cohesiveness, which has otherwise been a concern in the literature on group therapy.¹³³ Concern with the group matrix has also not spread beyond Karterud to other theorists of mentalization.¹³⁴

The limited explicit concern with group culture, cohesion, and hierarchy by MBT therapists has been criticized by some patients, who feel that sometimes the 'elephant in the room' is not raised by therapists when this appears to relate to a group-level rather than individual-level phenomenon.¹³⁵ Other patients report their impression that group MBT is asking them to mentalize the group, suggesting that this task is felt by some as expected or demanded by group MBT, even if this not a goal of the developers of MBT or the therapists delivering

¹²⁸ Bateman, A. W. and Fonagy, P. (2004). *Psychotherapy for Borderline Personality Disorders: Mentalization Based Treatment*, Oxford: Oxford University Press, pp. 176, 263.

¹²⁹ Potthoff and Moini-Afchari recommend a 'binocular' approach in which mentalization-based interventions are combined with attention to group-level processes. In this recommendation, it is assumed that mentalization-based group-therapy is concerned solely with individual minds, and offers little relevant to reflection or intervention with group-level processes. Potthoff, P. and Moini-Afchari, U. (2014). 'Mentalization-Based Treatment in Groups—A Paradigm Shift or Old Wine in New Skin?'. *Group Analysis*, 47(1): 3–16.

¹³⁰ Karterud, S. (2015). *Mentalization-Based Group Therapy (MBT-G)*, Oxford: Oxford University Press, p. 70.

¹³¹ Foulkes, S. H. (1971). 'The Group as Matrix of the Individual's Mental Life', in *Selected Papers: Psychoanalysis and Group Analysis*, London: Karnac Books, pp. 223–234. See also Karterud, S., Folmo, E., and Kongerslev, M. T. (2019). 'Personality and the Group Matrix'. *Group Analysis*, 52(4): 503–519.

¹³² Karterud, S. (2015). *Mentalization-Based Group Therapy (MBT-G)*, Oxford: Oxford University Press, p. 90.

¹³³ Hornsey, M. J., Dwyer, L., and Oei, T. P. (2007). 'Beyond Cohesiveness: Reconceptualizing the Link between Group Processes and Outcomes in Group Psychotherapy'. *Small Group Research*, 38(5): 567–592.

¹³⁴ One exception: there is an extremely passing reference to the matrix concept in Fonagy, P., and Adsheed, G. (2012). 'How Mentalisation Changes the Mind'. *Advances in Psychiatric Treatment*, 18(5): 353–362.

¹³⁵ Morken, K. T. E., Binder, P. E., Arefjord, N. M., and Karterud, S. W. (2019). 'Mentalization-Based Treatment from the Patients' Perspective—What Ingredients do they Emphasize?'. *Frontiers in Psychology*, 10. In a recent chapter, Bateman and colleagues have emphasized the importance of 'active management of process of the group'. However, no more detail on what this entails is provided than i) managing the level of group arousal, and ii) 'interventions that cultivate and maintain cohesion early in the group treatment'. Bateman, A., Kongerslev, M. and Bo, S. (2019). 'Group Therapy for Adults and Adolescents', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp.117–133, pp. 130–131. However, elsewhere in the same book, Bateman and colleagues observe that 'group work stimulates a hierarchical process within a peer group, which can be harnessed in vivo by the clinicians working with the group to explore participants' sensitivity to hierarchy and authority and the mentalisation distortions that ensue'. Bateman, A., Fonagy, P. and Campbell, C. (2019). 'Antisocial Personality Disorder in Community and Prison Settings', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 335–349, p. 340.

it.¹³⁶ In fact, occasional remarks by Fonagy and colleagues have suggested ambivalence about what relationship they wish patients to have with the group. For instance, Fonagy and Adshead encourage therapists to ‘engender attachment bonds between members of the group, and with the group as a whole’.¹³⁷ But this is not a recommendation that has recurred in subsequent papers.

Mentalizing cultures

In 2019, Asen, Campbell, and Fonagy published a chapter in the *Handbook of Mentalising in Mental Health Practice* specifically focused on social systems. As well as the culmination of previous partial attempts to address the topic, the choice to focus a chapter on social systems might be placed in the context of two other trends. First, Jurist has observed that mentalization theory must surely muster some response to the destruction of communal supports and dehumanizing government policies. These throw into relief the extent to which environments capable of sustaining mentalization and epistemic trust are dependent on societal infrastructures and collective processes, and we have seen how these can be broken or dismantled.¹³⁸

Direct attention to social systems in the *Handbook of Mentalising in Mental Health Practice* might also be placed in the broader context of increased attention to this topic over the past decade from across various disciplines in the social sciences and humanities. For instance, work in neuroscience has documented that the activation of brain regions associated with mentalization, such as the medial prefrontal cortex, temporo-parietal junction, and precuneus, do not distinguish individual from group targets. The activation appears to be the same for participants reflecting on the potential ‘intentions’ of an individual *or* an organization.¹³⁹ In fact, people can attribute mental states (thoughts and feelings) to a group even if they attribute those mental states to none of the group’s individual members, and vice versa.¹⁴⁰ In 2011, Fonagy praised the growing development of attention to group-level processes in organizations and societies within psychoanalytic theory.¹⁴¹ The need for renewed attention to group-level processes has similarly been urged in recent years by Twemlow, who has explicitly reappraised his work with Fonagy in terms of an intervention to improve community-level processes in the schools.¹⁴² However, Twemlow drops the mentalizing

¹³⁶ See e.g. Arefjord, N., Morken, K., and Lossius, K. (2019). ‘Comorbid Substance Use Disorder and Personality Disorder’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 403–416: ‘It makes me dizzy sometimes, trying to mentalise myself, the group, and my family, it is too much for me ... [with a smile]’ (p. 414).

¹³⁷ Fonagy, P. and Adshead, G. (2012). ‘How mentalisation changes the mind’. *Advances in Psychiatric Treatment*, 18(5): 353–362, p. 355. See also Esposito, G., Karterud, S., and Freda, M. F. (2019). ‘Mentalizing Underachievement in Group Counseling: Analyzing the Relationship between Members’ Reflective Functioning and Counselors’ Interventions’. *Psychological Services*, Early View: ‘The crucial interventions were exploratory questions on mental states addressed to the individual or to the group and defense interpretations.’

¹³⁸ Jurist, E. (2018). *Minding Emotions: Cultivating Mentalisation in Psychotherapy*, New York: Guilford Press.

¹³⁹ Jenkins, A. C., Dodell-Feder, D., Saxe, R., and Knobe, J. (2014). ‘The Neural Bases of Directed and Spontaneous Mental State Attributions to Group Agents’. *PLoS One*, 9(8): e105341. For a parallel in phenomenology, see Salice, A. and Taipale, J. (2015). ‘Group-Directed Empathy: A Phenomenological Account’. *Journal of Phenomenological Psychology*, 46(2): 163–184.

¹⁴⁰ Jenkins, A. C., Dodell-Feder, D., Saxe, R., and Knobe, J. (2014). ‘The Neural Bases of Directed and Spontaneous Mental State Attributions to Group Agents’. *PLoS One*, 9(8): e105341.

¹⁴¹ Fonagy, P. (2011). Back cover endorsement for E. Hopper and H. Weinberg (eds), *The Social Unconscious in Persons, Groups and Societies*, Volume 1: Mainly Theory, London: Karnac Books.

¹⁴² E.g. Koh, E. and Twemlow, S. W. (2017). ‘Towards a Psychoanalytic Concept of Community (III): A Proposal’. *International Journal of Applied Psychoanalytic Studies*, 14(4): 261–272, p. 264.

framework in doing so, presumably following Fonagy's classical assumption that only individuals can be the subject or target of mentalizing.

In their 2019 chapter, Asen and colleagues reflected that the long-standing focus on individual mentalizing since Fonagy's work in the late 1980s was itself shaped by cultural values, which trained attention on individual responses to challenges. They acknowledged that, in their culture, an excessive causal role is attributed to individual agency compared with the contribution of contextual and collective factors.¹⁴³ By contrast, 'in other cultures, the central unit is not the body or "self" of the individual, but the community and especially the family'.¹⁴⁴ For individuals to be judged not to be mentalizing because they are concerned with the dynamics or cultural memory of communal units rather than individuals would seem ethnocentric.¹⁴⁵ Asen and colleagues accept that mentalisation-based interventions may risk exactly such ethnocentrism to the extent that they treat a patient's attention to the beliefs and affects of collective units as either outright non-mentalizing, or as a distraction from the real work of attending to the beliefs and affects of individuals. And to the degree that a therapeutic modality is institutionally ethnocentric, it is likely to flounder, or at least face serious obstacles, in attempts to generate epistemic trust.¹⁴⁶

The same would be true for other communal-level processes, such as social antagonisms, wealth inequalities, and structural discrimination. The 'distributed cognition' possible in friendships, considered earlier in the chapter, and in group problem solving would likewise be excluded as a legitimate target of mentalizing.¹⁴⁷ As conventionally formulated, if patients were to raise such matters during MBT, technically this would be considered not

¹⁴³ In a recent conference presentation, Fonagy has taken this further, reviewing research on the propensity to underestimate the importance of social factors. Fonagy, P. (2019). 'The Future Prospects of Mentalization-Based Therapies', 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November: 'Our view of the person overemphasizes individual agency.' Fonagy discussed a paper by Haslam and colleagues, which showed that, while social integration and social support are among the most potent predictors of life expectancy, participants rated them the least potent from 11 options. Haslam, S. A., McMahon, C., Cruwys, T., Haslam, C., Jetten, J., and Steffens, N. K. (2018). 'Social Cure, what Social Cure? The Propensity to Underestimate the Importance of Social Factors for Health' *Social Science & Medicine*, 198: 14–21.

¹⁴⁴ Asen, E., Campbell, C., and Fonagy, P. (2019). 'Social Systems: Beyond the Microcosm of the Individual and Family', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, p. 233. Fonagy, P. and Campbell, C. (2019). 'Supporting the Social Triad. A Commentary on Keeping Culture in Mind: A Systematic Review and Initial Conceptualization of Mentalizing from a Cross-Cultural Perspective', *Clinical Psychology: Science and Practice*, 26(4): e12305: 'Coming from an attachment theory background, we have emphasized the individual's experience of being inadequately mentalized as the root cause of an incapacity to trust. This may have been inappropriate and was probably the consequence of our own cultural bias.' Jurist has explicitly argued that the boundaries of mentalizing should be large enough to include an individual's attention to, and reflection on, cultural memory. Jurist, E. (2018). *Minding Emotions: Cultivating Mentalisation in Psychotherapy*, New York: Guilford Press, p. 141.

¹⁴⁵ The argument that the mentalizing or and by collective units needs to be considered to avoid an ethnocentric theory has been developed in Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast/>; Slides. Accessed at: <https://www.borderlinepersonalitydisorder.org/wp-content/uploads/2019/02/Podcast-2.24.19-compressed.pdf>. A parallel argument has recently been made by Pereira and Debbané. Drawing on Foucault, they have claimed that mentalization-based therapists should not just facilitate mentalization of i) self and ii) other, but also on iii) the cultural and institutional context of the therapeutic work itself. Without this sociological reflective functioning, they worry that therapists risk becoming insensitive to, or overburdened by, the power-relations that structure the therapeutic encounter, and as a result lose track of the basis and qualities of the patient's mental states. Pereira, J. G. and Debbané, M. (2018). 'An Integrative-Relational Approach in Schizophrenia: From Philosophical Principles to Mentalization-Based Practice', in I. Hipólito, J. Gonçalves, and J. G. Pereira (eds), *Schizophrenia and Common Sense*, New York: Springer, pp. 193–207.

¹⁴⁶ This latter point is implicit in the 2019 chapter, but was made explicitly by Fonagy in Fonagy, P. and Wampold, B. (2016). 'Psychotherapy Debate with Peter Fonagy and Bruce Wampold at the Nordic Conference on Mental Health'. Accessed at: <https://www.youtube.com/watch?v=U5fhfAZnduU>.

¹⁴⁷ See e.g. Laughlin, P. R. (2011). *Group Problem Solving*, Princeton: Princeton University Press.

mentalizing, and the therapist would be enjoined to redirect the patient's attention towards the thoughts and feelings of individuals, specifically.¹⁴⁸ MBT therapists may even see civic disobedience and political activism, perhaps correctly, as sustained by the sense of urgency from psychic equivalence. Individual therapists may adapt their practice, but technically the modality would suggest redirecting patients away from this kind of thinking. Critics such as Parker have criticized Fonagy and colleagues, alleging that MBT is inherently normalizing, with 'patients effectively told how they could or should feel about key events in their lives or about their relationships with others.'¹⁴⁹ This criticism is certainly too strong. Nonetheless, there is a more subtle normativity in MBT. The modality aims to get patients to engagement in mentalizing, but with mentalizing defined in politically and culturally circumscribed ways. Asen and colleagues appear to acknowledge this as a problem in recognizing that the modality was shaped by cultural values that directed focus to individual responses to challenges.

The Asen and colleagues chapter has three ramifications for the fundamental conceptualization of mentalization:

1. It implies recognition of the potential for shared wishes and plans, for 'we' or 'they'—not merely 'I'—to be the agent of action, to possess thoughts, and to sustain feelings. For 'us' to establish an institution, buy a flat, or perform an opera is not merely the aggregate of our efforts, purchases, or performances, but an irreducibly collective undertaking and commitment.¹⁵⁰ The 'irreducibly collective mode of cognition called the we-mode' has been highlighted explicitly by Fonagy in recent workshops and conference presentations,¹⁵¹ and experimental support has been offered by Csibra and colleagues, who have shown that the results of judgements made on the basis of 'we-intentions' differ from those occurring on the basis of I-intentions.¹⁵² Even in the

¹⁴⁸ Lyotard would call this a 'differend': the patterned and potentially wholly inadvertent subtraction of relevance from particular forms of social suffering, through a practice that promises acknowledgement—but does so on the basis of limited terms. Lyotard, J. (1988). *The Differend: Phases in Dispute*, trans. G. Van den Abbeele, Minneapolis, MN: University of Minneapolis Press. Our point here is regarding the general logic of MBT as set out in written texts, rather than how Fonagy or Bateman would individually practise as clinicians. In fact, there appears to be only one instance where Fonagy and colleagues mention patients discussing political views within MBT, and the nature of the case makes it difficult to extrapolate: Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'Many MBT-ASPD groups begin with dialogue seemingly in pseudo-philosophical-political mode about how awful the system is, how no one can be trusted, how the police are corrupt, and so on. In essence, this is pretend mode. Contrary to the usual exhortation about pretend mode (challenge it and do not allow it to become embedded), it is important to allow the participants to do this at the beginning of a group to give them a sense of unity' (p. 383); 'The aim of the group is to encourage participants to identify their current feeling in the group rather than to express bitterness about external organisations, and to increase their recognition of the way context influences their current feeling' (p. 406). See also Aiello, G. and Pariante, C. M. (2013). 'Citizen, Interrupted: The 2011 English Riots from a Psychosocial Perspective'. *Epidemiology and Psychiatric Sciences*, 22(1): 75–79. Here the London riots are interpreted as reflecting a lack of mentalizing on the part of the rioters. The authors acknowledge that 'those arrested during the riots mainly came from deprived areas and had the poorest educational backgrounds' (p. 75), but do not apparently regard this as having played a causal role.

¹⁴⁹ Parker, I. (2019). *Psychoanalysis, Clinic and Context*, London: Routledge, p. 183.

¹⁵⁰ Jankovic, M. and Ludwig, K. (eds), (2017). *The Routledge Handbook of Collective Intentionality*, London: Routledge; Salmela, M. and Nagatsu, M. (2017). 'How Does it Really Feel to Act Together? Shared Emotions and the Phenomenology of We-Agency'. *Phenomenology and the Cognitive Sciences*, 16(3): 449–470.

¹⁵¹ 'The Future Prospects of Mentalization Based Therapies', 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November, discussing Gallotti, M. and Frith, C. D. (2013). 'Social Cognition in the We-Mode'. *Trends in Cognitive Sciences*, 17(4): 160–165. Also Fonagy, P. (2020). 'Trust and Interdependence: Lessons from the Study of Human Attachments'. UCL Minds Event, virtual talk, 25 June.

¹⁵² Török, G., Pomiechowska, B., Csibra, G. and Sebanz, N. (2019). 'Rationality in Joint Action: Maximizing Coefficiency in Coordination'. *Psychological Science*, 30(6): 930–941.

archetypal scene of parental reflective function, ‘we-intentions’ may be relevant: for instance, when a parent displays marked mirroring and ostensive cues relating to same-sex gender roles or a collective family culture or project (‘yes, darling, that’s how we do it’). Conversely, if ‘we’ or ‘they’ can sustain ‘we-intentions’, then ‘we’ (or ‘they’) can also be subject to mentalizing. Asen and colleagues appear to be moving towards the treatment of social or cultural settings as legitimate targets of mentalizing—to the degree that these units have collectively held beliefs and feelings.

2. A second implication is that intentions, thoughts, and feelings can be shaped by the qualities and projects of collective units. This might include their distributions of social status or material resources, lines of coercion or solidarity, and collectively held values or affective atmospheres.¹⁵³ Consideration of the role of these units in facilitating or hindering mentalizing may have earlier been regarded as either non-mentalizing or as a distraction from mentalizing. In 2012, Asen and Fonagy had explicitly drawn the comparison that ‘whereas systemic practitioners might seek explanations in the individual’s context, whether it is their family, social or cultural setting ... an MBT-F therapist may, under certain circumstances, view this as a non-mentalizing stance to adopt’.¹⁵⁴ The comments by Asen, Campbell, and Fonagy in 2019 suggest a wish to clarify the circumstances in which attention to social or cultural settings should not be regarded as non-mentalizing. This aligns with Fonagy’s growing conviction that ‘poor mental health is not simply in our individual minds. It is part of a social condition.’¹⁵⁵
3. Third, Asen and colleagues appear now to suggest that different collective units, including societies and institutions, may differentially facilitate or hinder different forms of mentalizing (e.g. of the self or of others) and of non-mentalizing (e.g. pretend mode or psychic equivalence). This suggests directions for elaborating the developmental model of mentalization and epistemic trust in the future to take into account cultural moderators and mediators. For instance, Aival-Naveh and colleagues have argued that ‘the suspension of epistemic mistrust may be dependent not only on accurate mentalizing and ostensive cueing from an attachment figure, but also on norms and social roles, most notably—filial piety and respect’.¹⁵⁶

Despite these advances, the 2019 chapter gives the impression of remaining suspended between two different models of social systems, without thorough integration. For instance,

¹⁵³ See Seyfert, R. (2012). ‘Beyond Personal Feelings and Collective Emotions: Toward a Theory of Social Affect’. *Theory, Culture & Society*, 29(6): 27–46; Salmela, M. (2012). ‘Shared emotions’. *Philosophical Explorations*, 15(1): 33–46; Anderson, B. (2014). *Encountering Affect: Capacities, Apparatuses, Conditions*, London: Routledge; Mercer, J. (2014). ‘Feeling Like a State: Social Emotion and Identity’. *International Theory*, 6(3): 515–535; Krueger, J. and Szanto, T. (2016). ‘Extended Emotions’. *Philosophy Compass*, 11(12): 863–878; Smith, E. R. and Mackie, D. M. (2016). ‘Group-Level Emotions’. *Current Opinion in Psychology*, 11: 15–19; Collins, S. (2018). ‘“The Government Should be Ashamed”: On the Possibility of Organisations’ Emotional Duties’. *Political Studies*, 66(4): 813–829.

¹⁵⁴ Asen, E. and Fonagy, P. (2012). ‘Mentalization-Based Therapeutic Interventions for Families’. *Journal of Family Therapy*, 34(4): 347–370, p. 351.

¹⁵⁵ Fonagy, P. (2019). ‘Mental Health is a Care We must Share’. *Guardian*, 13 October. Accessed at: <https://www.theguardian.com/society/2019/oct/13/isolation-not-social-media-cause-teenager-mental-ill-health>.

¹⁵⁶ Aival-Naveh, E., Rothschild-Yakar, L., and Kurman, J. (2019). ‘Keeping Culture in Mind: A Systematic Review and Initial Conceptualization of Mentalizing from a Cross-Cultural Perspective’. *Clinical Psychology: Science and Practice*, 26(4): 25: ‘different aspects of mentalizing may be more important depending upon culture. Consequently, it can be expected that different cultures will exhibit different mentalizing profiles ... we cannot simply assume that people across cultures have a similar ability, or need, to perceive internal mental states.’ See also Fonagy, P. and Campbell, C. (2019). ‘Supporting the Social Triad. A Commentary on Keeping Culture in Mind: A Systematic Review and Initial Conceptualization of Mentalizing from a Cross-Cultural Perspective’. *Clinical Psychology: Science and Practice*, 26(4): e12305.

even on a single page (p. 233), the authors present two different definitions of the term ‘culture’:

Culture can be defined as a dynamic system of explicit and implicit rules established by groups in order to ensure their survival, involving attitudes, values, beliefs, traditions, customs, norms and behaviours ... Culture could also be described as a system of knowledge shared by a relatively large group of people, passed on from one generation to the next.¹⁵⁷

It would appear that the authors have two different objects in mind, though they give them both the same name, expressing—apparently without realizing—a wider divergence in social scientific practice between two uses of the term ‘culture’.¹⁵⁸ A first sense is ‘anthropological culture’, taken to mean a collective form of life that shapes the perceptions and intentions of its members, with particular priority given to shared myths, presuppositions, and taken-for-granted practices. Anthropological culture also encompasses the role of social institutions and forms of self-perception in sustaining these myths, presuppositions, and practices. A second sense is ‘cognitive informational culture’, taken to mean the net information held by members of a group.¹⁵⁹ In the former, the social system is conceived as having a degree of autonomy from its members, with group-level dynamics and attributes.¹⁶⁰ In the latter, the social system is conceived as the sum of the information held by its members.¹⁶¹

To illustrate the difference: when a set of tools or concepts depend upon extensive specialist knowledge for their use, sustaining epistemic vigilance regarding their appropriate application hinders their spread into anthropological culture but not their ascension into cognitive informational culture. An example is the Strange Situation as a technical research

¹⁵⁷ Asen, E., Campbell, C. and Fonagy, P. (2019). ‘Social Systems: Beyond the Microcosm of the Individual and Family’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, p. 233.

¹⁵⁸ Duncan, J. S. and Duncan, N. G. (2004). ‘Culture Unbound’. *Environment and Planning A*, 36(3): 391–403; Fornäs, J. (2017). *Defending Culture: Conceptual Foundations and Contemporary Debate*, London: Palgrave. Though not immediately relevant to this discussion, it may be helpful to note that a third sense of ‘culture’ also operates within social scientific discourse, distinct from anthropological culture and cognitive informational culture. This is culture as ‘refinement’.

¹⁵⁹ In psychoanalytic terms, anthropological culture can be described as prioritizing rows C (myths) and D (preconceptions) on Bion’s grid—elements that do not form discrete and determinate thoughts, and may reside as much in the community as within the minds of individuals. Cognitive informational culture can be described as prioritizing rows E (conception) and F (concept)—elements that form relatively more refined, individual thoughts. Bion, W. R. (1963). *Elements of Psychoanalysis*, London: Karnac Books. To draw a distinction between anthropological culture and cognitive informational culture is not to imply that they are not compatible. For instance, Vygotsky held that anthropological culture organizes the kinds of tasks a child faces and the kinds of cognitive informational tools provided or withheld for responding to those tasks. Luria, A. R. (1979). *The Making of Mind: A Personal Account of Soviet Psychology*, trans. Michael Cole and Sheila Cole, Cambridge, MA: Harvard University Press, p. 44.

¹⁶⁰ This model of culture seems to be presumed when Allison and Campbell write about ‘cultural barriers to considering the perspectives of children and young people’. Examples given by the authors include ‘the stigma around mental health problems, which challenge the limits of our capacity for understanding and can evoke fearful responses, and the systemic tendency for the political process to neglect children because they do not have the vote’. Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre, p. 11. Cf. Bonanno, G. A., Romero, S. A., and Klein, S. I. (2015). ‘The Temporal Elements of Psychological Resilience: An Integrative Framework for the Study of Individuals, Families, and Communities’. *Psychological Inquiry*, 26(2): 139–169.

¹⁶¹ The orientation of Fonagy and colleagues towards this latter definition of culture is certainly influenced by Gergely. However, an earlier and background influence may be Winnicott who, after acknowledging that the term has multiple meanings, stated that ‘in using the word culture I am thinking of the inherited tradition. I am thinking of something that is in the common pool of humanity, into which individuals and groups of people may contribute, and from which we may all draw if we have somewhere to put what we find’. Winnicott, D. W. (1967). ‘The Location of Cultural Experience’. *The International Journal of Psychoanalysis*, 48: 368–372, p. 370.

tool, constrained in its use by expertise and resource demands. When a set of tools or concepts are circulated in a 'light' commodified version that requires little or no specialist knowledge, sustaining epistemic credulity regarding their appropriate application hinders their use within cognitive informational culture but precisely facilitates their spread within anthropological culture. An illustration is the idea of 'attachment', which has seen widespread take-up, but with poor alignment with how the concept was used in Bowlby's technical writings.¹⁶² Another example might be the concept of 'mentalization' itself, which can have diverging roles in attempts to transfer technical information (cognitive informational culture) and in the communal processes that operate through shared myths and presuppositions operating within psychological research and practice (anthropological culture).

The two uses of the term 'culture' are reflected, and not reconciled, in uses of the concept of epistemic trust by Asen, Campbell, and Fonagy in thinking about social systems:

1. In some cases, the term 'social system' is used to refer to anthropological culture and explore collective characteristics that facilitate or hinder individuals and groups from learning from experience.¹⁶³ This is clearest in the account by Asen and colleagues of the work of Bevington and Fuggle on creating organizational structures facilitative of mentalization-based integrative treatment. Indeed, there is friendly criticism of Fonagy's writings in Bevington and Fuggle's wish for 'a shift away from the hope or expectation of creating individual mentalising' in their chapter of the *Handbook*.¹⁶⁴
2. In other cases, however, and perhaps predominantly, the term 'social system' is used by Asen, Campbell, and Fonagy in a way that makes it seem to represent the sum of individual qualities that help or hinder learning from experience. This learning might be at a group or institutional level or at an individual level, but, with culture interpreted as cognitive information held by members of a group, mentalizing remains solely the preserve of individuals. Collective mentalizing capacity falls into the background.

The two concepts of culture are not reconciled, and at times produce significant incoherence in the chapter. For instance, the authors seem unsure about whether the welfare state i) improves the aggregate epistemic trust and mentalizing of citizens and/or ii) reduces violence through shaping the collective norms, priorities, and concerns of institutions and the wider society.¹⁶⁵

In contrast to the relatively poor coherence in their discussion of epistemic trust, in their 2019 chapter, Asen, Campbell, and Fonagy are more explicit than in earlier work about the existence of collective-level mentalizing processes. They even argue that it is possible to be 'mentalised by relevant social systems', not just individuals, and that individuals can 'attempt

¹⁶² Duschinsky, R. (2020). *Cornerstones of Attachment Research*, Oxford: Oxford University Press.

¹⁶³ See also Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9: Compromise of the processes of social learning 'is a systemic failure of communication that may characterize a family, the members of a social group such as a gang, a social subculture, or indeed an entire culture'.

¹⁶⁴ Bevington, D. and Fuggle, P. (2019). 'AMBIT: Engaging the Client and Community of Minds', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 211–228, p. 221.

¹⁶⁵ The same ambiguity is present in Lorenzini, N., Campbell, C. and Fonagy, P. (2019). 'Mentalisation and its Role in Processing Trauma', in Bernd Huppertz (ed.), *Approaches to Psychic Trauma: Theory and Practice*, Lanham, MD: Rowman & Littlefield, pp. 403–422.

to mentalise social systems', not just one another.¹⁶⁶ This is an important statement, re-adjusting the fundamental boundaries of mentalization theory. Teams, groups, organizations, societies are now to be treated as able to give and receive recognition; this makes them 'partial persons' in terms of their capacity to mentalize and be mentalized, on the basis of beliefs and affective atmosphere implicated in joint actions, preferences, and duties.¹⁶⁷ One case is the work of Fonagy and colleagues themselves. Over decades, this research collective have themselves been distinctive in their remarkable capacity to incorporate learning from other specific collectives and research traditions within the contested space of academic knowledge.¹⁶⁸

In recent presentations, Fonagy has proposed that this turn to attention to social systems and communities represents a 'Copernican revolution'.¹⁶⁹ The experience of being mentalized by social systems, more than mentalization by individual caregivers, may be the primary basis for the potential for epistemic trust, via secure attachment to and identification with the community. He has suggested that 'the infant is responding to the sensitivity of a network towards their sense of agency, how much the network is paying attention to them, not a single individual. That is what fosters trust.' As such, 'we ignore too much the importance of the infant recognizing the specificity of the other as an instance of the general other. This particular other is part of the network of "my group of people, my community" to which I belong, and represents my identity.'¹⁷⁰ A sense of belonging to the community facilitates epistemic trust in that community. Fonagy has therefore urged a shift in the study of child development to 'assess the caregiver network's sensitive responsiveness', and test the contribution of this collective capacity to the development of epistemic trust, and conversely to the development of personality disorders.¹⁷¹ He expressed concern that accounts of mentalizing

¹⁶⁶ Asen, E., Campbell, C. and Fonagy, P. (2019). 'Social Systems: Beyond the Microcosm of the Individual and Family', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, p. 235.

¹⁶⁷ However, it is unclear whether Asen and colleagues regard the capacity for individuals to be 'mentalised by relevant social systems' as implying some concept of group mind or 'group experience'; it would depend on what was meant, and the kind of social system in question. Cf. Fonagy, P. (2017). 'The Big Kahuna: Countries with the Highest Incidence of BPD'. Accessed at: <https://www.youtube.com/watch?v=lcVSMuxApAY>: 'To me, mentalising is not something that exists within an individual. Not even something that exists between a mother and a child. Not even something that exists between a family. It is something that a community of human beings owns'.

¹⁶⁸ Cf. Wilholt, T. (2016). 'Collaborative Research, Scientific Communities, and the Social Diffusion of Trustworthiness', in M. S. Brady and M. Fricker (eds), *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*, Oxford: Oxford University Press, pp. 218–233.

¹⁶⁹ Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>; Slides. Accessed at: <https://www.borderlinepersonalitydisorder.org/wp-content/uploads/2019/02/Podcast-2.24.19-compressed.pdf>. Fonagy, P. (2019). 'The Future Prospects of Mentalization Based Therapies', 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November.

¹⁷⁰ Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>; Slides. Accessed at: <https://www.borderlinepersonalitydisorder.org/wp-content/uploads/2019/02/Podcast-2.24.19-compressed.pdf>.

Fonagy cites Weisner: 'The question that is important for many, if not most, parents and communities is not, "Is [this individual] child 'securely attached'?", but rather, "How can I ensure that my child knows whom to trust and how to share appropriate social connections to others?" Weisner, T. S. (2015). 'The Socialization of Trust: Plural Caregiving and Diverse Pathways in Human Development across Cultures', in H. Otto and H. Keller (eds), *Different Faces of Attachment*, Cambridge: Cambridge University Press, pp. 263–277, p. 263.

¹⁷¹ Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>; Slides. Accessed at: <https://www.borderlinepersonalitydisorder.org/wp-content/uploads/2019/02/Podcast-2.24.19-compressed.pdf>. See also Dagan, O. and Sagi-Schwartz, A. (2018). 'Early Attachment Network with Mother and Father: An Unsettled Issue'. *Child Development Perspectives*, 12(2): 115–121. The contribution of in-group/out-group dynamics to mentalizing and epistemic trust had earlier been discussed briefly in Luyten, P. and Fonagy, P. (2015). 'The Neurobiology of Mentalizing'. *Personality Disorders*, 6: 366–379.

and epistemic trust have over-emphasized the importance of family interactions, and underestimated the importance of modern sociological processes that isolate families and hinder the functioning and integrity of communities, undermining their capacity to show sensitivity to individuals' emergent sense of personal agency. Rather than treat 'resilience' as equivalent to epistemic trust (see Chapter 7), Fonagy has now proposed that we 'conceptualize "resilience" less as a quality held by the child, and more as an outcome of the quality of the social network surrounding the child and the child's capacity to access that network'.¹⁷² In the context of the COVID-19 lockdown, Fonagy has argued that this situation has especially highlighted just how limited the nuclear family is, and how important the wider social network, for meeting children's needs.¹⁷³

In their 2019 chapter, Asen and colleagues develop a framework for thinking about the harms of collective non-mentalizing. They give the example of organizations in which 'psychic equivalence becomes institutional—it characterizes the function of a social system'.¹⁷⁴ The paradigmatic example, for them, is George Orwell's dystopian novel *Nineteen Eighty-Four*, in which thoughts and feelings are treated as real political threats, and therefore must be controlled, with a single viewpoint defended rigidly and even violently. However, Asen and colleagues observe that many more familiar social systems—schools, health services, political infrastructures—develop related characteristics as psychic equivalence becomes institutional.

Similarly, pretend mode and teleological mode can likewise become institutionalized in the culture and even the protocols of organizations.¹⁷⁵ In pretend mode, especially in the form of hypermentalizing, within the institution there may be 'endless "communication" and searching but it is destined to yield no change'. Meanwhile, there may be all kinds of awful behaviour contrary to the organization's stated values, which are not registered due to the dominance of pretend mode.¹⁷⁶ In institutionalized teleological mode, only concrete outcomes

¹⁷² Fonagy, P. (2018). 'Meeting the Mental Health Needs of Looked-After Children and Care Leavers', in Marc Bush (ed.) *Addressing Adversity: Prioritising Adversity and Trauma-Informed Care for Children and Young People in England*, London: Young Minds, pp. 170–179, p. 176. Accessed at: <http://www.instituteofhealthequity.org/resources-reports/addressing-adversity-prioritising-adversity-and-trauma-informed-care-for-children-and-young-people-in-england>.

¹⁷³ Fonagy, P. (2020). 'Trust and Interdependence: Lessons from the Study of Human Attachments'. UCL Minds Event, virtual talk, 25 June. See also Lassri, D. and Desatnik, A. (2020). 'Losing and regaining reflective functioning in the times of COVID-19: Clinical risks and opportunities from a mentalizing approach'. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(S1): S38.

¹⁷⁴ Asen, E., Campbell, C. and Fonagy, P. (2019). 'Social Systems: Beyond the Microcosm of the Individual and Family', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalizing in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, p. 231.

¹⁷⁵ An example is given in passing by Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'Intriguingly, the justice system has, by and large, adopted the same teleological stance. The logic of "justice being seen to be done" is acceptable to prisoners and guards in equal measure' (p. 64). The potential for externalization of the alien self to become institutionalized has been a familiar concern in psychodynamic organizational theory, under the rubric of projective identification. See e.g. Hinshelwood, R. D. and Chiesa, M. (eds), (2002). *Organisations, Anxieties and Defences: Towards a Psychoanalytic Social Psychology*. New York: Wiley-Blackwell. The series editors were Fonagy and Hepworth.

¹⁷⁶ The characteristics of institutionalized pretend mode are not detailed in the 2019 chapter, but appear in an earlier conference presentation with substantial overlap in content: Fonagy, P. (2012). 'Mentalization and Attachment: The Implication for Community Based Therapies'. Paper presented at the Community of Communities 10th Annual Forum, March, London. Accessed at: <https://www.slideshare.net/raffaelebarone/mentalization-and-attachment-the-implication-for-community-based-therapies>: 'Social systems that create fear and hyperactivate attachment can destroy thinking capacity and force the system back to pre-mentalistic modes of social thinking. Such social systems can be self-reinforcing and therefore highly stable in their instability. They undermine the very social mechanism that could alter their character: human collaboration (negotiation and creativity)'. Institutionalized pretend mode is also described in Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press: 'People

matter. For instance, New Public Management in the UK is a system oriented by teleological mode, in which the efforts of workers are not to be treated as real or meaningful unless they have demonstrable and successful consequences that can be audited and monitored by the system.¹⁷⁷ This is certainly the institutionalized culture of many contemporary health and social care services.¹⁷⁸ The institutionalization of teleological mode can also make organizations intensely concerned with perceived threats to their reputation, perceived as having catastrophic potential consequences.¹⁷⁹

Since the 1990s, Fonagy has from time to time offered his personal impression of a ‘dehumanisation, the stripping of humanity, that is occurring at all levels of culture’ in social systems.¹⁸⁰ This reflects major social and economic shifts, as well as the pervasive social exposure associated with new media, which in interaction have contributed to an increased recognition of mental illness and likely also a real rise in certain mental health issues among young people. Fonagy and colleagues have speculated that the increase in BPD may reflect a response to fragmented, dehumanizing social systems in the context of modernity. Fonagy has offered the examples of China and Russia as illustrating the societal conditions for the production of BPD, the prevalence of which has risen vertiginously in line with social fragmentation.¹⁸¹ However, he sees such particular cases as intensified forms of a sociological process characteristic of modernity more generally:

experience their thoughts and feelings as having no consequence for others, leading ultimately to the experience of an empty and meaningless social existence: this is a social system that is operating in the pretend mode. Selfishness and extreme egocentricism then emerge out of the unreality of anything other than one’s own thoughts and feelings’ (p. 451).

¹⁷⁷ New Public Management is comprised of seven doctrines: 1) Hands-on and active management; 2) Explicit standards and measures of performance; 3) Emphasis on measuring the quality of outputs; 4) Breaking up state services so that they can be placed in competition; 5) Promoting competition between services; 6) Importing techniques for controlling workers from the private sector; 7) Disciplining costs and promotion of cost-effectiveness. McLaughlin, K., Osborne, S. P., and Ferlie, E. (eds), (2002). *New Public Management: Current Trends and Future Prospects*, London: Routledge. An early use of mentalization theory to think about the implications of these doctrines for an institution’s capacity to acknowledge staff thoughts and feelings was attempted by Siltala, J. (2013). ‘New Public Management: The Evidence-Based Worst Practice?’. *Administration & Society*, 45(4): 468–493.

¹⁷⁸ Bevington, D., Fuggle, P., Cracknell, L. and Fonagy, P. (2017). *Adaptive Mentalisation-Based Integrative Treatment: A Guide for Teams to Develop Systems of Care*, Oxford: Oxford University Press.

¹⁷⁹ It should be qualified that the institutionalization of non-mentalizing is not, in itself, a problem. Consider the court system: there is extensive pretend mode in all the court ritual, both symbolizing and masking that the institution is backed in the final instance by the violence of the state; this helps set the frame for psychic equivalence within which the court is felt as holding authority; the result is the capacity to treat behaviour and evidence in teleological mode—for instance, with the payment of damages treated as assuaging personal insult or injury. Each of these has its place for the rule of law, but also produces serious and structural obstacles to mentalizing in this environment.

¹⁸⁰ E.g. Fonagy, P. (1999). ‘Interview with Peter Fonagy’, in S. M. Stein and J. Stein (eds), *Psychotherapy in Practice: A Life in the Mind*, Oxford: Butterworth Heinemann, pp. 77–98: ‘An interest in the human mind is becoming less important as part of the professional work of a number of key groups, including doctors ... This is a major danger, a danger in how we handle ourselves as people and how society works. I don’t think it is the doctors’ fault, nor is it due to the advancement of science—it is a cultural thing. It is the dehumanization, the stripping of humanity, that is occurring at all levels of culture. If you look at television, for example, you will see people being portrayed as behaving like machines. The “cops and robbers” kind of stories where people are shot or shoot without intention, without desire and without belief. They also die without intention ... I think this is mirrored in medical education, which is allowing itself to focus on technical advantages while de-emphasizing the more human aspects of patients ... I’ve been teaching medical students for seventeen years and I’ve noticed a change ... I don’t know what’s causing it, nor what the main driving force behind it is, but I think it is culture-wide’ (p. 84). See also Fonagy, P. (1999). ‘Male Perpetrators of Violence against Women: An Attachment Theory Perspective’. *Journal of Applied Psychoanalytic Studies*, 1(1): 7–27: ‘Society has relinquished its caretaking function, demolished its institutions for supporting emotional development’ (p. 23).

¹⁸¹ Fonagy, P. (2019). ‘Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience’. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>. At the same time, the increased use of non-mentalizing prompted by fragmented, dehumanizing social systems in turn can reinforce them: ‘Social systems that create fear and hyperactivate attachment can destroy

The anomie of modern life—that is, a lack of social connectiveness leading to dysregulation—described by Durkheim, and connected by other authors with the conditions that might account for national variations in BPD, can be read as a description of a systemic collapse of epistemic trust.¹⁸²

The comparison with Durkheim is revealing. What is shifting in modernity in general, and which is revealed most starkly in China and Russia, is not simply a breakdown of the experience of the claims and experience of others as dependable, generalizable, or relevant. It is a *systematic* collapse—i.e. a collapse at a systemic level, rather than only something occurring for individuals. For Durkheim, the shift away from solidarity based on strong social ties, in the context of modernity, does not alter the fact that individuals fundamentally remain shaped by their cultural and economic ecology, which channels relatively shared forms of moral and social perception and interpretation based on shared practices.¹⁸³ This cultural and economic ecology will facilitate or undermine our experience of the claims and experience of others as dependable, generalizable, or relevant. For Durkheim, the most important aspect of this ecology was economic inequality, which hinders the achievement of solidarity between individuals.¹⁸⁴ In the terms of Fonagy and colleagues, it hinders collective mentalizing capacity.¹⁸⁵

In countries like the UK, Fonagy has speculated that the availability of mental health services and improvements may be mitigating the rise in BPD that would otherwise have been expected in the fragmented, dehumanizing social systems that have taken hold in so many sectors over recent decades.¹⁸⁶ The epidemiology of BPD in part reflects its social construction by infrastructures of psychological classification. However, additionally, Fonagy has suggested that its epidemiology to an extent reflects the relative presence of anomie. Within societies, for example, Fonagy has proposed that the greater incidence of BPD among women may reflect their dehumanization and objectification, in which full subjectivity is tacitly ascribed more to men than to women in countless ways even while equality has ostensibly been achieved.¹⁸⁷ As one case, Fonagy has drawn attention to the way that,

thinking capacity and force the system back to pre-mentalistic modes of social thinking. Such social systems can be self-reinforcing and therefore highly stable in their instability. They undermine the very social mechanism that could alter their character: human collaboration (negotiation and creativity)' Fonagy, P. (2012). 'Mentalization and Attachment: The Implication for Community Based Therapies'. Paper presented at the Community of Communities 10th Annual Forum, March, London. Accessed at: <https://www.slideshare.net/raffaelebarone/mentalization-and-attachment-the-implication-for-community-based-therapies>.

¹⁸² Fonagy, P., Luyten, P., Allison, E., and Campbell, C. (2017). 'What We have Changed our Minds About: Part 2. Borderline Personality Disorder, Epistemic Trust and the Developmental Significance of Social Communication'. *Borderline Personality Disorder and Emotion Dysregulation*, 4(1): 9. For instance, in the case of parents: 'the relative social isolation in which modern parents engage in caregiving makes it much harder to access another trusted person/mind to help think about stressful situations in order to regulate affect and regain balanced mentalizing', Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre.

¹⁸³ Rawls, A. W. (2012). 'Durkheim's Theory of Modernity: Self-Regulating Practices as Constitutive Orders of Social and Moral Facts'. *Journal of Classical Sociology*, 12(3–4): 479–512.

¹⁸⁴ Durkheim, E. (1984 [1893]). *The Division of Labour in Society*, trans. W. D. Halls, Book 3. London: Macmillan.

¹⁸⁵ For an interpretation of solidarity in terms of mentalizing, see Luyten, P., Nijssens, L., Fonagy, P., and Mayes, L. C. (2017). 'Parental Reflective Functioning: Theory, Research, and Clinical Applications'. *The Psychoanalytic Study of The Child*, 70(1): 174–199, p. 179.

¹⁸⁶ Fonagy, P. (2019). 'Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience'. GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>.

¹⁸⁷ Fonagy, P. (1999). 'Epistemological and Methodological Background', in Fonagy, P. (ed.), *An Open Door Review of Outcome and Process Studies in Psychoanalysis*, London: International Psychoanalytic Association, pp. 8–85, p. 50. See also Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable*

despite discourses of gender equality, in practice ‘we have made mothers inappropriately and uniquely responsible’ for children and child care.¹⁸⁸ In a recent newspaper article, he has highlighted that there is greater pressure on women students in higher education as they do exams and think about job market prospects, because expectations of success and equality are mismatched with an unequal economic and social reality.¹⁸⁹

In 2015, Fonagy and Allison characterized mentalization theory as aligned with liberation politics, proposing that it ‘overlaps almost completely with queer theory’ in seeking and defending ‘the singular truth of the individual’s experience’ against ‘publicly accepted forms of knowledge.’¹⁹⁰ To claim almost complete overlap with queer theory was unpersuasive, because at the time Fonagy and colleagues had no account of publicly accepted forms of knowledge,¹⁹¹ or of the formation of gender identities.¹⁹² With the 2019 chapter by Asen, Campbell, and Fonagy, earlier remarks about the potential harms caused by social systems are integrated with the theory of mentalizing in the conceptualization of non-mentalizing modes ‘becoming institutional’. This opens intriguing possibilities for thinking about forms of oppression like sexism and homophobia in terms of problems in collective mentalizing capacity, and the specific roles of pretend mode, psychic equivalence, and teleological mode. First, however, Fonagy and colleagues will need to develop a much more robust account of dominance, oppression, and social consecration to make sense of:

- who is regarded as having the right to offer ostensive cues;
- who is attributed real and socially pertinent mental states;

Development, London: Anna Freud Centre: ‘racial discrimination is a powerful predictor of general psychopathology (Clark, Anderson, Clark, & Williams, 1999), with the most powerful associations observed for depression and anxiety (Banks, Kohn-Wood, & Spencer, 2006; English, Lambert, Evans, & Zonderman, 2014) and conduct problems ... Being at the receiving end of essentialist homogenizing beliefs creates generic vulnerability through disrupting effective communication and the potential of adaptation and adjustment through learning. Within this framework, prejudices of all kinds, but particularly prejudice based on ethnic divisions, can generate psychological problems by disrupting the natural flow of information within social networks.’

¹⁸⁸ Fonagy, P. (2018). ‘Peter Fonagy: Combating a Mental Health Crisis’. Accessed at: <https://www.goldmansachs.com/insights/talks-at-gs/peter-fonagy.html>.

¹⁸⁹ Fonagy, P. cited in Thomson, A. and Sylvester, R. (2020). ‘Panic and Anxiety after Education is Plunged into Limbo’. *The Times*, 31 March.

¹⁹⁰ Fonagy, P. and Allison, E. (2015). ‘A Scientific Theory of Homosexuality for Psychoanalysis’, in A. Lemma and P. E. Lynch (eds), *Sexualities: Contemporary Psychoanalytic Perspectives*, Hove, UK: Routledge, pp. 125–137, p. 130–131.

¹⁹¹ The claim to almost complete overlap with queer theory was also unpersuasive given the lack of sustained social criticism by Fonagy and colleagues. One step towards making this claim more real has been that the Anna Freud Centre joined London Pride in 2019. Fonagy gave a statement that: ‘It is profoundly wrong that people we know and love fear rejection from their friends and families, from their colleagues and workplaces, feel vulnerable to abuse and fear violence simply because of their sexual identity. Being accepted is a human need. It is not a privilege to have that need met, but a basic right, one from which all other opportunities flow. When we can all assert our individuality and flourish, and when we have the respect we need to feel proud of who we are, we will have a fairer society. It’s in recognition of this aspiration that the Anna Freud Centre is marching in celebration and solidarity with Pride.’ Accessed at: <https://www.annafreud.org/insights/news/2019/07/anna-freud-centre-staff-students-and-champions-march-to-mark-london-pride/>. The Centre issued guidance from one of their Young Champions on how mental health services can better serve as safe spaces for people who identify as LGBTQ+: accessed at: <https://www.annafreud.org/insights/blogs/2019/07/how-mental-health-services-can-present-themselves-as-safe-spaces-for-people-who-identify-within-the-lgbtqplus-community/>. Recently, the Centre have also responded critically to the judicial review of the Gender Identity Development Service, advocating greater acknowledgement of the capacity of trans and gender-questioning young people to self-determination. <https://www.annafreud.org/insights/news/2020/10/statement-from-the-anna-freud-centre-on-the-gender-clinic-judicial-review/>.

¹⁹² Straker, G. and Winship, J. (2011). ‘The Dangers of the Universal: A Critique of Fonagy and Target’s Theory of Sexual Enjoyment’. *Studies in Gender and Sexuality*, 12(4): 288–302.

- whose signals are subject to epistemic vigilance;
- the costs or punishments associated with deviating from socially established conventions for ostension, for the recognition of mental states, and for the enactment of epistemic vigilance.¹⁹³

Such an account has precisely been one of the major achievements of queer theory, and has been fundamentally absent in the work of Fonagy and colleagues. It is not false to regard queer theory as pitching the singular truth of the individual's experience against publicly accepted forms of knowledge; this is one important element.¹⁹⁴ More important to queer theory, however, is the role of publicly accepted forms of knowledge and practice in organizing and situating who is considered to have experience and the psychological and epistemic standing of that experience. To take an example, Ahmed has examined the gendered relations of power that can construct epistemic vigilance enacted by women as a distraction, enforcing epistemic trust and framing the women themselves as 'spoilsports' or 'killjoys'.¹⁹⁵

An important obstacle for Fonagy and colleagues over the years in conceptualizing dominance, oppression, and social consecration has been that the scene that has served as their lens on mentalization, and more recently on epistemic trust, has been the caregiver–infant relationship. Not that they have ignored that this relationship may be abusive; far from it. But the scene of recognition or misrecognition is of simple wishes and needs. Though an infant is dependent on the power of the caregiver, Fonagy and colleagues do not consider this scene as one in which there is any contest for dominance, any interest or manipulation on the part of the caregiver, and no necessity for the caregiver to justify his or her authority. At least by 12 months, these features are present, albeit perhaps not necessarily salient in child–caregiver interactions; it would have been quite different if Fonagy and colleagues had narrated the scene of parental reflective function with the child as a true preschooler. It was always a peculiar mismatch that in expositions of their theory in the 1990s and early 2000s by Fonagy and colleagues, the ages 2–3 were given priority for the development of mentalizing, but in accounts of parental reflective function the child seemed to have the characteristics of an infant, displaying simple intentions, but not battling to jump into the pond, wanting to wear their sibling's clothes, or refusing to put down the attractively packaged product grabbed from the aisle in the supermarket.¹⁹⁶ A scene of parental reflective function without conflict, deception, or cultural mediation has been retained over subsequent decades, even as the prioritization of preschool for the development of mentalizing has been abandoned by Fonagy and colleagues, and even as their focus has shifted from marked mirroring to ostensive cues and the development of epistemic trust.

¹⁹³ Sedgwick, E. K. (2008). *Epistemology of the Closet*, Berkeley, California: University of California Press; Butler, J. (2005). *Giving an Account of Oneself*, New York: Fordham University Press; Berlant, L. (ed.) (2019). *Reading Sedgwick*, Durham, NC: Duke University Press.

¹⁹⁴ Huffer, L. (2012). 'Foucault and Sedgwick: The Repressive Hypothesis Revisited'. *Foucault Studies*, 14: 20–40.

¹⁹⁵ Ahmed, S. (2010). *The Promise of Happiness*, Durham, NC: Duke University Press; Ahmed, S. (2017). *Living a Feminist Life*, Durham, NC: Duke University Press.

¹⁹⁶ The problem is actually acknowledged by Bateman and colleagues in a paper from 2007, though not subsequently rectified. The scene of thinking about parental reflective function remained infancy: Bateman, A. W., Ryle, A., Fonagy, P., and Kerr, I. B. (2007). 'Psychotherapy for Borderline Personality Disorder: Mentalization Based Therapy and Cognitive Analytic Therapy Compared'. *International Review of Psychiatry*, 19(1): 51–62: 'the problem with the concept of contingent and marked mirroring is that it appears to offer a one-sided and reductive account of the maternal–infant interaction ... The mentalizing perspective is a dynamic developmental view where the respective capacities of child and parent, and therefore the nature of their contribution to interactions, change as the child matures. While the mentalizing approach considers that contingent mirroring may be the key contribution of the parent in the first year of life, this gives way to more complex interaction' (p. 55).

Queer theorists would ask: what has been the implication of the scenes in which Fonagy and colleagues predominantly imagine mentalizing taking place? These scenes are, above all, dyadic: the analytic clinic and parental reflective functioning. They would point to a certain epistemic bias in the image of what counts as mentalizing, because the way it has been imagined has tended to filter out the collective conditions, social antagonisms, and practical workaday phenomena beyond the dyad that might otherwise be recognized as facilitating or obstructing mentalization. One example, for instance, is the neglect of Fonagy and colleagues of the role of habit in facilitating or obstructing mentalizing.¹⁹⁷ Another is their neglect of social oppression. It is notable that a quality of the four key ideas that Fonagy and colleagues drew from Gergely—marked mirroring, teleological mode, ostensive cues, epistemic trust is that each offers a way of considering the relationship between individual theory of mind and embeddedness in sociality. This is part of what makes them such powerful concepts and heuristics. However, each concept simultaneously limits the scope of acknowledgement of sociality. So, for instance, marked mirroring is irreducibly dyadic and social, but the dyad is removed from cultural context, adversities, and the pragmatic challenges of everyday life.¹⁹⁸ Fonagy and colleagues know, of course, that these ideas are pragmatic simplifications. However, the problem has been that, in incorporating them, Fonagy and colleagues tended over the decades to lose track of what has been lost.¹⁹⁹ For instance, in characterizing epistemic trust in terms of an individual open to learning from other individuals or from society, Fonagy and colleagues bracket the way that cognitive informational culture is invested by relations of power and exclusion, which influence what is supported to become knowledge. As black feminist epistemologists such as Hill Collins and Dotson have argued, the tools and skills that sustain cognitive informational culture are differentially accessible and, when accessed, differentially suited for use by dominant groups.²⁰⁰ The capacity to recognize oneself as a learner, or for others to recognize us as someone from whom they could learn, are likewise structured by *intersecting* forms of authority and prestige, invisibility, and denigration, which apply to both individuals and communities. As is the capacity to recognize mental states in others and in oneself. These processes are missed as, following

¹⁹⁷ The role of habit in facilitating or hindering awareness of thoughts and feelings has been a classic concern of phenomenology since Husserl, and has continued in the work of queer theorists developing the phenomenological tradition e.g. Ahmed, S. (2006). *Queer Phenomenology: Orientations, Objects, Others*, Durham, NC: Duke University Press.

¹⁹⁸ An aligned criticism has been offered by Køster, A. (2017). 'Mentalization, Embodiment, and Narrative: Critical Comments on the Social Ontology of Mentalization Theory'. *Theory & Psychology*, 27(4): 458–476: 'in spite of a strong interactive focus, it remains fundamentally rooted in a Cartesian ontology, overlooking embodied, expressive, enactive and sociocultural dimensions of social cognition' (p. 460); 'Bernhard Waldenfels (2015) distinction between "frontal" vs. "lateral" variants of sociality. Whereas frontal sociality refers the type of "face-to-face interaction" which completely dominates the framework of MT, lateral sociality designates the equally prominent dimension that we find ourselves "side-by-side" in a shared relation to a cultural world of practices, norms and objects etc. which are saturated with social reference and meaning. Whereas the early attachment context might supply the child with a basic understanding of the rhythm and structure of this lateral dimension, it hardly seems persuasive to claim that these forms of social understanding exclusively or even primarily owe their ontogenesis to the early infant/caretaker interaction. Rather, these norms are learned from actual engagement in these social practices in their proper contexts' (p. 467).

¹⁹⁹ Similar concerns have been raised by Shaw, C., Lo, C., Lanceley, A., Hales, S., and Rodin, G. (2019). 'The Assessment of Mentalization: Measures for the Patient, the Therapist and the Interaction'. *Journal of Contemporary Psychotherapy*, 50(4). It may be noted that the concepts were also incorporated by Fonagy and colleagues into the developmental model quite early, on the basis of one or two empirical studies generally conducted by Gergely—that is to say, before much is known about factors that might clarify the roles of context and culture. That said, marked mirroring and teleological mode were introduced 20 years ago now, but moderators have still barely been considered.

²⁰⁰ Dotson, K. (2014). 'Conceptualizing Epistemic Oppression'. *Social Epistemology*, 28(2): 115–138; Dotson, K. (2017). 'Theorizing Jane Crow, Theorizing Unknowability'. *Social Epistemology*, 31(5): 417–430; Hill Collins, P. (2019). *Intersectionality as Critical Social Theory*, Durham, NC: Duke University Press, Chapter 4.

Gergely, Fonagy and colleagues give 'cognitive informational culture' precedence over 'anthropological culture' in their thinking about learning.

With anthropological culture more in view, queer theory draws from Althusser and Foucault in imagining a different scene of recognition and misrecognition to that of Fonagy and colleagues. Here, the encounter is not between child and caregiver but between individual and institution, or representative of that institution.²⁰¹ This fundamentally alters the stakes of questions about ostensive cues, the recognition of mental states, and the deployment of epistemic vigilance.²⁰² Even the scene of marked mirroring and ostensive cues in early childhood appears in a different light, because the way in which the caregiver comes to offer acknowledgement of the child's mental states, intention, and individuality is already contingent on forms of symbolic and material stratification, and a wider economic and cultural ecology. So, for instance, a caregiver's recognition of the child's playful exploration (see Chapter 5) will already be embedded within gendered systems of power, shaping what the child comes to recognize as their own curiosity and excitement, and the manner in which they feel they can dare to venture out into the world and learn from experience.²⁰³ As we saw in Chapter 6, for Fonagy and colleagues the self is partly imaginary, a confabulation weaving together diverse sources of information. This already offers an important qualification to any idea of 'reflective' function as simple, transparent knowledge. Yet, a further complexity left aside by Fonagy and colleagues, at least until recently, is the way that self-representation is always already saturated by a wider economic and cultural ecology ('the symbolic' in Lacanian terms), which not only prompts and informs the self-representation but also the integration of its elements, including the identification of intentions excluded from the self-representation to form part of the 'alien self'.²⁰⁴ Seemingly moving in this direction, in

²⁰¹ Althusser, L. (1971). 'Ideology and Ideological State Apparatuses', in *Lenin and Philosophy*, trans. Ben Brewster, New York: Monthly Review Press; Foucault, M. (1982). 'The Subject and Power', *Critical Inquiry*, 8(4): 777–795. Curiously, Winnicott has moments when he makes links between the parent–child scene and the wider scene of cultural subjectivation. However, in these cases, society is always mediated by the mother. His accounts are also almost always positive, with these two elements acting in potential concert: 'The mother's confidence in her husband or in the support that she will get, if she calls out, from local society, perhaps from the policeman, makes it possible for the child to explore crudely destructive activities which relate to movement in general, and also more specifically destruction that has to do with the fantasy that accumulated round the hate. In this way (because of environmental security, mother supported by father, etc.) the child becomes able to do a very complex thing, that is to say, integrate all his destructive impulses in with the loving ones' Winnicott, D. W. ([1967] 1986). 'Delinquency as a Sign of Hope', in Clare Winnicott, Ray Shepherd, and Madeleine Davis (eds), *Home is Where We Start From: Essays by a Psychoanalyst*, New York: Norton, pp. 90–100, p. 94.

²⁰² What is 'shown' to the infant in ostension—for instance, the caregiver's desire for the child to pay attention to particular information—is already embedded in forms of symbolic and material stratification, and a wider economic and cultural ecology. Marcia Cavell has criticized Fonagy directly on these grounds. Cavell, M. (2006). *Becoming a Subject: Reflections in Philosophy and Psychoanalysis*, Oxford: Oxford University Press, p. 69. Likewise, see Straker, G. and Winship, J. (2011). 'The Dangers of the Universal: A Critique of Fonagy and Target's Theory of Sexual Enjoyment', *Studies in Gender and Sexuality*, 12(4): 288–302. For relevant work in queer and feminist theory, see e.g. Butler, J. (2004). *Undoing Gender*. London: Routledge; Benjamin, J. (2005). 'What Mothers and Babies Need: The Maternal Third and its Presence in Clinical Work', in S. Brown (ed.), *What do Mothers Want?* New York: Analytic Press, pp. 37–54. Object relations theory has been an important influence on the attention of queer theorists to infancy. However, attention to this site has also helped contest the classical liberal characterization, at least since Locke, of child–caregiver relationships as somehow outside of the forms of oppression that structure wider society.

²⁰³ Young, I. M. (2005). *On Female Body Experience: 'Throwing Like a Girl' and Other Essays*, Oxford: Oxford University Press.

²⁰⁴ Lacan, J. (1977). *Écrits: A Selection*, London: Tavistock; Gasché, R. (1986). *The Tain of the Mirror: Derrida and the Philosophy of Reflection*, Cambridge, MA: Harvard University Press. That these were complexities knowingly left aside is shown by the fact that, in their 2003 book, Fonagy and Hepworth offer a summary precisely of Lacan's critique of psychoanalysts who over-focus on the imaginary aspects of the mirror stage at the expense of the symbolic. Fonagy, P. and Target, M. (2003). *Psychoanalytic Theories: Perspectives from Developmental Psychopathology*, London: Whurr Publications, p. 17.

conference presentations in 2019, Fonagy has argued that the ‘infant recognises the specific other as an instance of the generic other characteristic of the community’, acknowledging the cultural situatedness of recognition within the child–caregiver dyad.²⁰⁵

A published consideration of recognition and misrecognition, not by the parent but by collective units, has also been offered recently by Fonagy and colleagues in thinking about the role of distorted intentions and/or mental states for individuals seeking to cope with social institutions that discount their needs, identities, and experiences. In the *Handbook of Mentalising in Mental Health Practice*, Fonagy, Campbell, and Allison characterized ‘distorted social cognition’ as entirely ‘necessary for people to be able to live with the vivid human capacity for social imagination and function in an inherently socially imperfect world.’²⁰⁶ Non-mentalizing modes might be used more by individuals with significant histories of adversity and oppression, as part of their adaptation to this adversity (see Chapter 7). However, Fonagy and colleagues contend, we all engage in distorted social cognition to the degree that we must learn to cope with a society and institutions ‘in which the minds of individuals are either discounted altogether or systematically described as “culpable” or “undeserving.” To varying degrees, we have each experienced distortions and damage to the imaginative capacity that underpins both mentalization and ethical vision as part of becoming people able ‘to tolerate the social system well enough to navigate it’. The institutionalization of non-mentalizing modes hinders our ‘capacity for both critique and creativity.’²⁰⁷

This account of the role of institutionalized non-mentalizing in mediating structural oppression and personal indifference is distinctive, and represents a potential advance in thinking about dehumanizing institutional arrangements. An example can be seen in the attention Fonagy has given to the ways in which collective non-mentalizing capacity may be facilitated by the affordances of new media technologies. Fonagy has characterized new media as a double-edged sword: currently contributing to social and psychological pressures on young people, but with significant potential to facilitate mentalization, self-care, and the strengthening of relationships with others.²⁰⁸ He has argued against a technological determinist argument in which the pervasiveness and the image-focused aspects of new media are necessarily contrary to mentalization. Instead, he has proposed that social

²⁰⁵ Fonagy, P. (2019). ‘The Future Prospects of Mentalization Based Therapies’, 5th International Congress of Mentalisation Based Treatments, Haarlem, The Netherlands, 22 November; Fonagy, P. (2019). ‘Why is it So Hard to Learn to Do Things Differently? On Not Being Able to Learn from Experience.’ GAP Call-In Series Podcast. Accessed at: <https://www.borderlinepersonalitydisorder.org/gap-call-in-series-podcast>.

²⁰⁶ Fonagy, P., Campbell, C. and Allison, E. (2019). ‘Therapeutic Models’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 169–180, p. 178. It is interesting that the term ‘social cognition’ is used here. As noted in Chapter 3, mentalization was introduced to characterize three of the four elements that Dunn had ascribed to social cognition. These were understanding of the i) origin, ii) location, and iii) functioning of mental states. Dunn had also considered children’s iv) understanding of social and cultural conventions. In referring to ‘social cognition’ here, there is tacit acknowledgement that distorted understanding of social and cultural conventions is also operational. In the Introduction to the *Handbook* (p. 4), Fonagy and Bateman indeed reiterate that mentalization is not coextensive with social cognition, but rather ‘is the aspect of social cognition that enables individuals to make sense of the behaviour of themselves and others’, without reference to social and cultural conventions.

²⁰⁷ Ibid.

²⁰⁸ Wallwark, E. (2016). ‘Thousands of Children Start School ‘Not Ready for the Classroom’, Parents’ Smartphone Habits to Blame’. Huffington Post, 09/05/2016. Accessed at: http://www.huffingtonpost.co.uk/entry/children-not-school-ready_uk_57308969e4b0ade291a2342e: ‘Professor Peter Fonagy, chief executive of the Anna Freud Centre agreed that smart phones don’t necessarily have to be a barrier to communication. “To the extent that smart phones stand in the way of that natural process they are indeed a problem,” he said. “However, what is stopping parents from using smart phones to have a conversation with children about how school went, what they have learnt, and for children [to] write about their reactions to new experiences? It may be perfectly possible to use smart phones to promote social relationships even though the tendency is for it to replace rather than enhance parent child interactions.”

media is good for circulating ostensive cues without the substance of recognition, and for stimuli for rapid, implicit rather than careful, explicit mentalizing.²⁰⁹ To date, their predominant effect has been to reduce sustained consideration of thoughts and feelings. However, he regards this as an effect of their embeddedness within social practices not built to promote mentalizing, rather than an inevitable effect of the technology itself.²¹⁰ He has given the example of his daughter, who was bullied at school, but who could use her phone to reach out to her parents rather than feeling left alone in the institutional context of the school.²¹¹

However, Fonagy and colleagues are yet to distinguish the social conditions that facilitate different forms of non-mentalizing—what would seem to be the greatest potential of a social systems interpretation of their theory.²¹² Furthermore, their hold on a social systems perspective remains precarious still. For instance, in the *Handbook of Mentalising in Mental Health Practice* chapter by Asen, Campbell, and Fonagy on social systems, the authors identify four characteristics of ‘a social system with a capacity for effective mentalising’.²¹³ The foremost characteristic of social systems with a capacity for effective mentalizing is that the system needs to be capable of retaining different perspectives rather than being ‘stuck in one point of view’. Second, the system ‘must be able to permit modifications of convention, at least on a temporary basis’. Though justification is not given, it would seem that this is regarded as facilitating the other characteristics of a social system with capacity for effective mentalizing. Third, a social system needs to have means of identifying and responding to the experiences and goals of its members, rather than assuming sufficient knowledge of these in

²⁰⁹ Duschinsky, R., Collver, J., and Carel, H. (2019). “‘Trust Comes from a Sense of Feeling One’s Self Understood by Another Mind’: An Interview with Peter Fonagy”. *Psychoanalytic Psychology*, 36(3): 224–227.

²¹⁰ E.g. Fonagy, P. (2018). ‘Peter Fonagy: Combating a Mental Health Crisis’. Accessed at: <https://www.goldmansachs.com/insights/talks-at-gs/peter-fonagy.html>: ‘With social media, and as kids have drawn back into their bedrooms and are using smartphones and are using a range of digital devices, what we have found is that they become less happy. Social media is not a replacement for social contact with real human beings. It is particularly when kids compare themselves ‘upwardly’, with models that are better. So for example body image, particularly among young women, is negatively related to screen time with Facebook and Instagram where there are photographs. Whereas looking just at text, your body image is unaffected.’

²¹¹ Interview with Peter Fonagy, cited in Rampton, M. (2019). ‘So your Kid Wants a Phone—Now What?’, *Huffington Post*, 18 October, https://www.huffingtonpost.co.uk/entry/what-age-give-child-a-phone_uk_5da6e5d0e4b062bdc1b3f29: “‘It’s the function of a phone that really matters, far more than the physical reality of having one,” he says. “It’s what a kid uses it for.” We can’t ignore the advantages of children having them. It’s great to be able to keep in touch with family and friends, says Fonagy, who stresses that phones aren’t inherently a bad thing. “After age nine or ten your main social network is other children, and using a phone to maintain that network seems, to me, good,” he says.’ He gives the example of when his daughter was bullied at school and was able, while sitting on her own and deeply upset, to reach out to her parents to share her distress.

²¹² In both editions of the *Handbook of Mentalising in Mental Health Practice*, Fonagy and colleagues highlight the potential relevance of their theory to conceptualizing war. However, in doing so, they do not advance much beyond Bion, who argued that one of the central objectives in warfare or other forms of sustained violence is to stop the other side thinking, and that lack of thinking helps perpetuate futile violence. Bion, W. R. (1940). ‘The War of Nerves’, in Miller, E. (ed.), *The Neuroses in War*, London: Macmillan, pp. 180–200. Differentiating the modes of non-mentalizing, this claim could be further specified to identify the role of pretend mode, psychic equivalence, and teleological mode (and externalization of the alien self) in offering false certainties or uncertainties, and blocking inquisitive stance about the minds of the enemy, allies, or one’s own forces. This line of thought may also help pursue a distinction, neglected in the work of Fonagy and colleagues, between mindless violence and mentalized violence—for instance, some forms of civil insurrection.

²¹³ These characteristics were already presented by Fonagy in conference presentations from 2012, and in Bateman and Fonagy 2016; they appear to reflect an older position less well integrated with recent thinking. Fonagy, P. (2012). ‘Mentalization and Attachment: The Implication for Community Based Therapies’. Paper presented at the Community of Communities 10th Annual Forum, March, London. Accessed at: <https://www.slideshare.net/raffaelebarone/mentalization-and-attachment-the-implication-for-community-based-therapies>; Bateman, A. W. and Fonagy, P. (2016). *Mentalization-Based Treatment for Personality Disorders: A Practical Guide* (2nd edn), Oxford: Oxford University Press, pp. 449–450.

advance. Finally, ‘the social system needs to emphasize the importance of individual “ownership” of, and responsibility for, specific behaviours and actions, rather than explaining individual behaviours or actions as the result of some external force.’²¹⁴

Yet there is a false opposition here. Why should awareness of causes of individual behaviour beyond the individual entail a denial or negation of individual responsibility? Would Asen and colleagues really treat the actions that the characters in *Nineteen Eighty-Four* are brought to as chosen, and as reflecting individual ownership and responsibility? As we read the book, especially the closing chapters, one of Orwell’s main concerns is with the way that feelings of personal responsibility can diverge terribly from the actual conditions for meaningful enactment of responsibility. As we saw in Chapter 6, Fonagy’s position has been increasingly qualified in recent years to balance two claims: where supports can make the illusion of personal agency somewhat meaningful and self-reinforcing, it can be a very helpful illusion because it becomes a kind of reality; where supports are unavailable to make the illusion of personal agency meaningful and self-reinforcing, the feeling of responsibility is persecutory and unhelpful. To take an illustration: if ‘we have made mothers inappropriately and uniquely responsible’ for children and child care,²¹⁵ the inappropriateness stems in part from the fact that the apparent uniqueness of this responsibility is part of a wider system of sexism, and in the fact that the unique responsibility is an illusion that masks the necessary contribution of external supports or lack of supports for the activities of caregiving.²¹⁶

The term ‘responsible’ needs to be considered carefully. Vincent, a philosopher, offers helpful distinctions, asking us to consider ‘the text of a notice that hangs in Café Doerak, my favourite bar in the Dutch city of Delft, “The management of this establishment is not responsible”’. Vincent points out that:

This notice is terribly ambiguous, and one might imagine two people engaged in a frustrating argument about it, simply because each understands it differently, though neither realizes that this is so. But the text of this notice could be helpfully re-written ...

CAPACITY: The management of this establishment are not (yet) psychologically mature.

VIRTUE: The management of this establishment are not dependable and might be reckless.

ROLE: The management of this establishment have no responsibilities towards its clients.

²¹⁴ Asen, E., Campbell, C. and Fonagy, P. (2019). ‘Social Systems: Beyond the Microcosm of the Individual and Family’, in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 229–243, pp. 234–235. By way of comparison, only the first and third of these characteristics are identified in the account of mentalizing organizations offered by Di Stefano, G., Piacentino, B., and Ruvo, G. (2017). ‘Mentalizing in Organizations: A Psychodynamic Model for an Understanding of Well-Being and Suffering in the Work Contexts’. *World Futures*, 73(4–5): 216–223. Di Stefano and colleagues identify some elements missing from the account of Asen, Campbell, and Fonagy. These include the role of the organization in regulating arousal, in treating thoughts and feelings as intelligible and of communal interest, and in offering an environment in which uncertainty and ambiguity can be tolerated. See also Koh, E. and Twemlow, S. W. (2018). ‘Towards a Psychoanalytic Concept of Community (IV): The Well-Functioning Community’. *International Journal of Applied Psychoanalytic Studies*, 15(1): 5–15.

²¹⁵ Fonagy, P. (2018). ‘Peter Fonagy: Combating a Mental Health Crisis’. Accessed at: <https://www.goldmansachs.com/insights/talks-at-gs/peter-fonagy.html>.

²¹⁶ Fonagy, P. and Target, M. (2002). ‘Fathers in Modern Psychoanalysis and in Society: The Role of Father and Child Development’, in J. Trowell and A. Etchegoyen (eds), *The Importance of Fathers: A Psychoanalytic Re-Evaluation*, New York: Brunner-Routledge, Taylor & Francis, pp. 45–66; Fonagy, P. and Higgitt, A. (2007). ‘The Early Social and Emotional Determinants of Inequalities in Health’, in G. Baruch, P. Fonagy, and D. Robins (eds), *Reaching the Hard to Reach: Evidence-Based Funding Priorities for Intervention and Research*, Chichester, UK: John Wiley & Sons, pp. 3–34. See also Duschinsky, R., Greco, M., and Solomon, J. (2015). ‘The Politics of Attachment: Lines of Flight with Bowlby, Deleuze and Guattari’. *Theory, Culture & Society*, 32(7–8): 173–195.

CAUSAL: The management didn't causally contribute to losses suffered on these premises.

OUTCOME: The management can't be blamed for whatever happens on these premises.

LIABILITY: The management won't pay for any losses suffered on these premises.²¹⁷

As has often occurred in Fonagy's use of the term 'responsibility' since the 1990s, Asen and colleagues seem to have conflated some of these meanings. They appear to have assumed that attention to the causal role of system-level factors is incompatible with an individual's sense of intentionality in their moral role, and thus with mentalizing.²¹⁸ As a result, the authors slip back into a stance in which attention to the beliefs and atmosphere of collective units as causal for individual behaviour is treated not simply as 'not mentalizing' but as non-mentalizing, destructive of individual mentalizing capacity.

Naturally, there may be situations in which collective mentalizing capacity and individual mentalizing operate at the expense of one another. But attention to the beliefs and atmosphere of collective units as causal for individual behaviour should not in itself be regarded as opposed to an individual's sense of intentionality. Jurist, among others, has identified this as, inadvertently, an ethnocentric, moralizing, and politically neoliberal stance.²¹⁹ It does not sit comfortably with the trend towards greater acknowledgement—and critique—by Fonagy and colleagues of the properties of the social system in shaping human capacities, including the capacity for mentalization and the viability of feelings of responsibility (see Chapter 6). For instance, elsewhere in recent work, Fonagy and colleagues have highlighted evidence that 'the prevalence of BPD can be relatively well predicted from the ratio of the average income of the richest 20% to that of the poorest 20% of the population. Hence, a general lack of social concern for equality may be directly related to the prevalence of BPD.'²²⁰ Such findings have led Allison and Campbell to argue that mental health difficulties in our society 'are a collective responsibility and ought to be a collective concern.'²²¹

Fonagy has recently argued against discourses of 'responsibility' in a society where young people are frequently made to feel powerless and encounter institutions and environments

²¹⁷ Vincent, S. (2011). '“A Structured Taxonomy of Responsibility Concepts”', in N. Vincent, I. van de Poel, and J. van den Hoven (eds), *Moral Responsibility: Beyond Free Will and Determinism*, New York: Springer, pp. 15–35, p. 26.

²¹⁸ E.g. Fonagy, P., Target, M., Steele, M., Steele, H., Leigh, T., Levinson, A., et al. (1997). 'Morality, Disruptive Behavior, Borderline Personality Disorder, Crime, and their Relationships to Security of Attachment', in L. Atkinson and K. J. Zucker (eds), *Attachment and Psychopathology*, New York: Guilford Press, pp. 223–274, pp. 256–257. Examining uses of the term 'responsibility' in Fonagy's work, it would seem that the problem was initiated by his use of the term in the 1990s to encompass the Kleinian problematic of guilt and reparation, in which accepting responsibility initiates the depressive position, acceptance of truth, and the basis for moral life with others. However, in Klein, this is responsibility for imagined attacks on the mother, rather than actions in reality. On this basis, taking inappropriate responsibility for faults that lie, in part, in environmental causes was regarded by Winnicott as no less pathological than the abdication of responsibility and the derogation of the environment. Winnicott, D. W. (1986). 'Berlin Walls', in C. Winnicott, R. Shepherd, and M. Davis (eds.), *Home is Where We Start From*, New York: W. W. Norton, pp. 221–227, p. 223.

²¹⁹ Jurist, E. (2018). *Minding Emotions: Cultivating Mentalisation in Psychotherapy*, New York: Guilford Press. Jurist encourages mentalization theory to become more 'cautious about using the term “emotion regulation”'. Emotion regulation has not emerged in a vacuum; it has come to the fore in a specific place and time, which valorises individual responsibility and is sceptical of communitarianism. Personally, I am wary of the moralising discourses that preach self-discipline, which can be associated with neoliberalism, and which obscure how dependent we are on others' (p. 48).

²²⁰ Luyten, P., Campbell, C., and Fonagy, P. (2019). 'Borderline Personality Disorder, Complex Trauma, and Problems with Self and Identity: A Social-Communicative Approach'. *Journal of Personality*, 88(1): 88–105.

²²¹ Allison, E. and Campbell, C. (2019). *Transforming Child Mental Health: Principles of Sustainable Development*, London: Anna Freud Centre. See also Fonagy, P. (2017). 'The Big Kahuna: Countries with the Highest Incidence of BPD'. Accessed at: <https://www.youtube.com/watch?v=lcVSMuxApaY>.

with inadequate support for enacting meaningful agency. Speaking to the Youth Select Committee:

I feel very strongly that, very differently from my own childhood, young people nowadays carry a far greater responsibility than I had to carry, in terms of my life course. My life course was far more set out and determined, and I didn't feel personally responsible for it. I actually had far more authority over my life than young people do these days. That combination of having to take responsibility without authority makes young people very vulnerable.²²²

Nonetheless, where supports can make the illusion of personal agency into a kind of reality, it can still be regarded as very helpful. "Acknowledgement that 'I am a person and I have agency'" is a perspective that supports the coherence and consistency of thoughts and feelings by supporting mentalizing and the capacity to make good use of social relationships.²²³

Behind the issue of responsibility in Asen, Campbell, and Fonagy's account of mentalizing culture lies ongoing ambiguities in the conceptualization of intention. This concept has an underpinning role across the work of Fonagy and colleagues. In part, it is linked to the contributions of Gergely. Intention is critical in distinguishing mentalization from teleological mode; and as the other object, besides mental states, that can be taken as the target of marked mirroring and ostensive cues. Beyond this, though, Fonagy and Allison have reconceptualized the psychoanalytic unconscious in terms of intentions externalized from the self-representation. However, there is an ambiguity to the term, contributing here to conceptual and normative problems with how Fonagy and colleagues in turn conceptualize responsibility. It is logical that mentalizing social systems would help people consider how behaviour and experience can be accounted for in terms of thoughts and feelings implicated in motivations and intentions. That is part of what mentalizing means. However, it is not logical that acceptance of individual responsibility excludes acknowledgement of external causal factors.²²⁴ This seems to depend on an essentialized notion of intention as individual will, relatively impervious to circumstances. It also runs contrary to statements made by

²²² Youth Select Committee (2017). 'Oral Evidence Taken before the Youth Select Committee on Friday 7 July. Accessed at: <http://www.byc.org.uk/wp-content/uploads/2017/10/Youth-Select-Committee-07.07.17-morning.pdf>. See also Chivers, T. (2016). '12 Things you Probably Didn't Know about Happiness'. Accessed at: <https://www.buzzfeed.com/tomchivers/12-scientific-facts-about-happiness>, citing an interview with Fonagy: 'the combination of responsibility and the absence of agency is a particularly toxic phenomenon'.

²²³ Asen, E. and Fonagy, P. (2017). 'Mentalizing Family Violence Part 1: Conceptual Framework'. *Family Process*, 56(1): 6–21: 'It would seem that the need for each individual to be recognized as a person in their own right is so powerful because, among other things, being recognized in this way is a precondition for the opening up of epistemic trust. To make knowledge resonate—to imbue it with epistemic trust—we need to feel that it is relevant to us, and this is linked to an acknowledgement that "I am a person and I have agency"' (p. 16).

²²⁴ Even Dennett, whose reification of intentionality was taken up by Fonagy and colleagues, acknowledges that it is a logical error: 'The intentional stance towards human beings, which is a precondition of any ascriptions of responsibility, may coexist with mechanistic explanations'. Dennett, D. (1981). *Brainstorms: Philosophical Essays on Mind and Psychology*, Cambridge, MA: MIT Press, p. 253. In a recent interview, Fonagy argues that mentalization allows one to act with 'agency' rather than to be a 'victim', at the mercy of what is evoked in us by our environment. BBC Radio 4 (2020). 'Peter Fonagy on a Revolution in Mental Health Care'. The Life Scientific Podcast, 28 January. Accessed at: <https://www.bbc.co.uk/programmes/m000dpj2>. It is wholly persuasive that difficulties in conceiving of and reconsidering the thoughts and feelings of others increase the likelihood that a person will feel that the actions of others are intended at their expense, contributing to a sense of being a victim. Yet, the opposition between mentalizing/agency and victimhood also seems to assume some of its appeal on the basis of wider cultural narratives that situate rationality, agency, and responsibility as opposed to weakness, victimhood, and influence by the environment. See e.g. Stringer, R. (2014). *Knowing Victims: Feminism, Agency and Victim Politics in Neoliberal Times*, London: Routledge; Johnstone, M. A. (2020). 'Plato on the Enslavement of Reason'. *Canadian Journal of Philosophy*, 50(3): 382–394.

Fonagy and colleagues elsewhere, on occasions when intention is considered more carefully, that state that one needs 'to feel responsible for and in control of one's own mental life as well as behaviour in the world, but also to assess fairly accurately what is outside one's control. Failures in the process of evaluation can lead at one extreme to a belief in one's own omnipotent responsibility'.²²⁵

Given that the concept of 'intention' is a key component in accounts of mentalization and epistemic trust, it would appear that slippage in use of the concept will be a continued hindrance for work to conceptualize the role of culture. On the one hand, when intention is treated as equivalent to individual will, this is likely to hinder attempts to make sense of the kinds and forms of intention involved in we-intentions, and in the ascription of intentionality to collective units. On the other hand, it will hinder attempts to understand the role of culture and context in shaping how we come to thoughts and feelings. For instance, reification of the idea of intention will contribute to the continued neglect of proto-thoughts and feelings, which play such a key role in shared myths, prohibitions, presuppositions, and taken-for-granted practices and artefacts ('anthropological culture'). Such issues may sustain the difficulties Fonagy and colleagues have had in integrating their concept of the primary unconscious with the rest of their thinking. Furthermore, given the turn in the most recent work of Fonagy and colleagues towards the pursuit of empirical trials of therapeutic interventions in the developing world,²²⁶ problems in their thinking about cultural differences may turn out to be particularly consequential.

²²⁵ Fonagy, P., Edgumbe, R., Target, M. and Miller, J. (1999). *Contemporary Psychodynamic Child Therapy: Theory and Technique*. London: The Anna Freud Centre and University College London, p. 54. See also Fonagy, P. (1998). 'Prevention, the Appropriate Target of Infant Psychotherapy'. *Infant Mental Health Journal*, 19(2): 124–150, p. 137.

²²⁶ For instance, a randomized-control trial of the use of play therapy for children in Ethiopia to accompany an emergency food intervention. Accessed at: <https://www.annafreud.org/what-we-do/research-policy/research-themes/testing-what-treatment-works-best/emotional-stimulation-in-the-context-of-emergency-food-intervention-in-the-treatment-of-malnourished-children-a-randomised-controlled-trial/>. Palmer, R. (2016). *Emotional Stimulation as an Addition to Therapeutic Food Intervention for Treatment of Young Children with Severe Acute Malnutrition in a Low-Income Country*. Unpublished doctoral thesis, London, University College London. Fonagy and colleagues are also pursuing a trial of group-based psychotherapy in Kenya and Lebanon. Accessed at: <https://www.nihr.ac.uk/news/research-funding-boost-for-mental-health-in-low-and-middle-income-countries/24900>.

Conclusions

Constructive appraisal

Throughout this book, we have attempted to offer reflections that have been friendly and constructive, looking to understand the underlying coherence of the theory as well as potential limitations. We acknowledge that it can be hard for a paradigm in psychological theory and practice to metabolize criticisms, because there is much invested in the current state of affairs, including ‘secondary gain’ from the fuzzy use of concepts and their lack of operationalization. As discussed in the Introduction, Sandler identified that there can be resistance to efforts to appraise or refine concepts that seem to work well enough for pragmatic purposes, and help hold a coalition of interested parties together, even if their use has had associated disadvantages:

The fact that they work well may in turn lead to an undue resistance to the progressive integration and modification of our concepts, so necessary for scientific development. This resistance can partly be overcome by the cultivation of a critical attitude towards our ideas, by discussion with colleagues, and by honest reading of the literature, but unless we are directly confronted with contradictions in our thinking, resistance to change due to secondary gains may prove too great.¹

An aspect of this book that we hope may help circumvent some of this resistance is that we have tried, whenever possible, to be constructive and specific with suggestions to accompany appraisal. A number of such reflections from across the book can be particularly highlighted, without intending to be exhaustive. In Chapter 3, we expressed concern about confusion in the use of the concept of disorganized attachment by Fonagy and colleagues. We proposed that processes such as i) fear of attachment figures, ii) dissociation, iii) conflicting feelings, iv) difficulties modulating feelings, and v) the absence of help-seeking in times of need should be addressed more explicitly, given that they have at times been lost or outright mistaken for one another when disorganized attachment has been invoked.² This may help resolve the ambivalence Fonagy has shown (see Chapter 7) regarding whether or not to include disorganized attachment in his developmental model of mentalizing and epistemic trust, and help retain attention to elements that otherwise fluctuate in and out of view, such as dissociation.

In Chapter 4, we identified both strengths and weaknesses in the ‘four poles’ model of mentalizing (inner/outer, automatic/controlled, affective/cognitive, self/other), which has

¹ Sandler, J. (1962). ‘Research in Psycho-Analysis—The Hampstead Index as an Instrument of Psycho-Analytic Research’. *International Journal of Psycho-Analysis*, 43: 287–291, p. 290.

² See also Duschinsky, R. and Solomon, J. (2017). ‘Infant Disorganized Attachment: Clarifying Levels of Analysis’. *Clinical Child Psychology and Psychiatry*, 22(4): 524–538; Solomon, J., Duschinsky, R., Bakum, L., and Schuengel, C. (2017). ‘Toward an Architecture of Attachment Disorganization: John Bowlby’s Published and Unpublished Reflections’. *Clinical Child Psychology and Psychiatry*, 22(4): 539–560.

been prominent in the theory and clinical guidance of Fonagy and colleagues since 2009. We offered a number of friendly suggestions of matters that might be addressed in future. First, the account of affective versus cognitive mentalizing is troubled by a lack of clarity regarding whether this refers to the *object* of mentalizing (e.g. mentalizing one's thoughts or feelings) or the *process* of mentalizing (e.g. mentalizing using one's thoughts or feelings). Second, the conditions that facilitate high-quality automatic mentalizing are left unaddressed. Third, 'automatic' mentalizing and 'outer' mentalizing are generally treated as non-mentalizing, leading to conceptual and terminological muddle between mentalizing and non-mentalizing. The treatment of automatic mentalizing as non-mentalizing also leads to the excluded middle of slow, semi-intentional 'feeling out', as in therapy. Fourth, even after a decade, no attempt has been made to operationalize the model. Perhaps in part this reflects difficulties with the constructs, but equally in turn the lack of attempts at operationalization has insulated the four poles model from scrutiny, empirical validation, or refinement.

The four poles model was introduced in 2009 to respond to commentators such as Jeremy Holmes, who had identified that the term 'mentalizing' was serving as something of a catch-all, impeding communication between and among researchers and clinicians, and hindering scientific discussion. In Chapter 4, we identified 28 different uses of the term 'mentalizing' over time by Fonagy and colleagues. In part, this represented the evolution of theory over time; in part, it represented a desire to encompass different processes within a single umbrella; in part, it expressed a degree of haziness in the theoretical conceptualization. Our reflections go beyond identification of these problems. On the basis of a comprehensive analysis of every usage of the term across the writings of Fonagy and colleagues, we synthesized a new definition of mentalizing:

The capacity to conceive of and make available for reconsideration the thoughts and feelings implicated in motivations and intentions in order to account for and explain the observable social behaviour and present and past perceptual experience of oneself and others.

This definition makes it easier to distinguish between full and specific forms of mentalizing, articulate the relationship between mentalizing and non-mentalizing, and specify the relationship between mentalization and mental states. It situates more exactly within one framework the dimensions that Fonagy and colleagues had described using the 'four poles', while resolving some of the difficulties with that account.

In Chapter 5, our new definition of mentalizing was used to address questions in the account of non-mentalizing offered by Fonagy and colleagues. The researchers had claimed that non-mentalizing represented a hijacking of mentalizing capacities, but without explaining how or why non-mentalizing deploys the very elements that comprise mentalizing. Our definition allowed us to address this question and specify more precisely what is being hijacked in non-mentalizing.

Pretend mode was characterized as using the capacity to conceive of thoughts and feelings implicated in motivations or intentions.

Hypermentalizing was characterized as a form of pretend mode using the capacity to conceive of thoughts and feelings implicated in motivations or intentions in order to account for and explain the observable social behaviour of others.

Psychic equivalence was characterized as using the capacity to account for and explain thoughts and feelings and observable social behaviour in terms of immediate experience.

Teleological mode was characterized as using the capacity to account for and explain observable social behaviour as implicated in the motivations or intentions of others and of oneself.

In articulating how non-mentalizing constitutes elements of mentalizing, this account helps specify when and how mentalization-based therapy (MBT) may contribute to harm to patients by entrenching non-mentalizing. An example, discussed in Chapter 8, is when MBT refines knowledge of the feelings of others, but where this is then used by the individual as a resource for relational aggression. A further advantage to our account of non-mentalising here is that it makes sense of an ambiguity in the work of Fonagy and colleagues: whether mentalizing is merely the absence of non-mentalizing, and vice versa. Our definition helps us resolve this question, because it highlights the centrality for mentalisation of the generative doubt that underpins inquisitive stance.

Chapter 6 drew together the threads of discussions of the 'self' across the writings of Fonagy and colleagues. Readers of the work of these researchers have generally been segmented, each audience having access to only a 'part object'. Fonagy and colleagues have also used the concept in different ways across these contexts. While perhaps inevitable for the sake of communication, there has been a price to pay for the theory—for instance, in a lack of clarity about how 'mental states' occur, and whether mentalizing one's own mental states is the same as or different from mentalizing others' mental states. Bringing together discussions of the self from across the writings of Fonagy and colleagues helped clarify the extent to which, for these researchers, the 'self' is a confabulation, as well as the role of two distinct kinds of unconscious processes in producing the content of this confabulation. We also raised the question about the status of 'externalization of the alien self', which at times shuttles between a description of experience and an explanation of experience. Certain versions of the concept would appear to represent a candidate for another form of non-mentalizing, a question Fonagy and colleagues may wish to address. However, we also note that the concept of 'externalization of the alien self', so prominent in their work in the 2000s, appears to have dropped out of favour in recent years, without explanation and while formally remaining on the books as part of the paradigm.

Chapter 7 identified that Fonagy and colleagues have increasingly sought to emphasize that non-mentalizing is an adaptive response to adversity, drawing ideas from attachment theory. However, this claim has not been supported by explanation of what is meant by adaptation. Using an unpublished interrogation of the concept of adaptation by Ainsworth, one of the founders of attachment theory, we identified three different meanings to the term: species-level natural selection, individual responsiveness to the environment, and individual thriving. Natural selection may have facilitated the development of mentalizing and forms of non-mentalizing as a repertoire of responses available to humans. And both mentalizing and non-mentalizing may—or may not—reflect responsiveness to contextual cues. As such, non-mentalizing is theorized by Fonagy and colleagues to contribute to short-term positive outcomes, especially in harsh environments. In turn, mentalizing is theorized to contribute to long-term positive outcomes, especially—or perhaps solely—in non-harsh environments. The question of whether contextual adversity moderates the effect of mentalizing and non-mentalizing on outcomes is an important one, and should certainly be addressed in future research.

Chapter 7 also offers some distinctions we hope will be constructive regarding the concept of epistemic vigilance, which has seen special prominence in the work of Fonagy and colleagues in recent years. A significant problem faced by their account is ambiguity between two concepts of epistemic vigilance: one in which the communication of others is subject to

scrutiny, and another in which the communication of others is outright distrusted. We recommended that the term ‘epistemic vigilance’ reverts back to the original usage of Sperber and colleagues, who introduced the concept, to mean a suite of cognitive mechanisms that help calibrate the credence assigned to the inferred experiences and claim others. This would be distinguished from ‘epistemic mistrust’ as a chronic state in which the thoughts (and perhaps feelings) of others as implicated in their motivations and intentions are not felt to be trustworthy, generalizable, or relevant to the self. Epistemic mistrust may be a transitory state; however, it can be stabilized by the non-mentalizing modes, and in turn contribute to non-mentalizing. The distinction between epistemic vigilance and mistrust is helpful in making sense of the claim of Fonagy and colleagues that epistemic vigilance is actually beneficial for mentalizing. Epistemic vigilance, but not epistemic mistrust, can support generative doubt and the potential for the reconsideration of the experiences of oneself and others, though the rigid and inflexible subjection of experiences to epistemic vigilance may contribute to hypermentalizing, rather than mentalization.

Chapter 9 offers reflections and suggestions relating to the account by Fonagy and colleagues of wider social systems extending beyond the individual. This is clearly an area of particular interest and theoretical development, and we hope that our distinctions can help facilitate future thinking and research in this area. One contribution was our efforts to excavate the concept of ‘collective mentalizing capacity’: the properties of the community that facilitate mentalization. In discussions of schools, family therapy, and group therapy, Fonagy and colleagues have offered partial acknowledgment of this process, and it is hoped that with attention drawn to collective mentalizing capacity the concept may see greater prominence, including further elaboration and ideally operationalization. The question of whether groups themselves can mentalize and non-mentalize, or only their individual members, has also been left in abeyance, and may be worth raising more focally in future because the stakes seem quite high. Relevant to this question, in Chapter 9, we identified an unacknowledged hinge in the use of the concept ‘culture’ by Fonagy and colleagues, and difficulties this was causing their account of social systems. In a first sense, they have been discussing ‘anthropological culture’, meaning a collective form of life that shapes the perceptions and intentions of its members, with particular priority given to shared myths, prohibitions, presuppositions, and taken-for-granted practices and artefacts. In a second sense, they have been discussing ‘cognitive informational culture’, the net information held by members of a group. This distinction is important because in focusing on cognitive information culture in their work on epistemic trust to date, the researchers have tended to neglect the role of anthropological culture, including the operation of oppression. It has also made it difficult for them to distinguish the social conditions that facilitate different forms of non-mentalizing, and that make individual responsibility feasible rather than a persecutory ideal.

Some remaining questions: mentalizing

One of the most important areas for future trials of MBT will be to examine whether improvements in mentalizing mediate the success of the modality, and how or whether improvement is moderated by social systems. Fonagy and colleagues have recently acknowledged this as a limitation in the current evidence base, and highlighted the importance of questions about working mechanisms.³ As such questions are pursued in future research, it would be good to

³ Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., and Fonagy, P. (2020). ‘Learning to Mentalize: A Mediational Approach for Caregivers and Therapists’. *Clinical Psychology: Science and Practice*, 27(3): e12334.

see the potential for iatrogenesis from MBT explicitly explored, ideally based on a means of measuring forms of non-mentalizing.

In terms of theory, one of the most important areas for future development will be to acknowledge and address limitations with the four poles model. In particular, the elements that appear to have particular value, such as mentalizing self versus mentalizing others, should be operationalized and/or incorporated into existing measures. Ensink's work adapting the reflective functioning scales has been a valuable step in this direction (see Chapter 4).⁴ The account of automatic mentalizing has been widely criticized by commentators, and might be revisited. In terms of future work, thinking about 'poles' of mentalizing might also address the difference between mentalizing a determinate mental state versus mentalizing inchoate or conflicted mental states. This might pick up on a difference relevant to clinical efforts to help patients enrich their mentalizing capacities. It may also help counter a tendency in the theorizing of Fonagy and colleagues to treat mental states as relatively determinate rather than conflicted or inchoate, though certainly they acknowledge that conflicted and inchoate mental states exist. In part as a result of this tendency, there are outstanding questions about the role of proto-thoughts and proto-feelings.⁵ It is not clear that facilitation of the capacity to conceive of the determinate mental states of others or oneself will have similar correlates to the capacity to conceive of the inchoate mental states of others or oneself. Clinical efforts to help the former capacity, when this is at the expense of the latter, may inadvertently contribute to teleological mode rather than mentalizing.

Another area where future work may be fruitful is in relation to anxiety. As we saw in Chapter 3, Fonagy and Target found that anxiety was a positive prognostic factor for children receiving therapy at the Anna Freud Centre. Yet there has been minimal subsequent attention to anxiety, despite the fact that it was a core concept for the Anna Freudian tradition of psychoanalysis inherited by Fonagy, such as in the work of Sandler.⁶ Even though dynamic interpersonal therapy is ostensibly an adaptation of MBT for people with depression or depression and anxiety, there is no sustained published discussion of how the treatment is presumed to address anxiety. Indeed, Fonagy appears to recommend cognitive-behavioural

⁴ Ensink, K. (2004). *Assessing Theory of Mind, Affective Understanding and Reflective Functioning in Primary School-Aged Children*. Unpublished doctoral dissertation, London: University College London; Ensink, K., Normandin, L., Target, M., Fonagy, P., Sabourin, S., and Berthelot, N. (2015). 'Mentalization in Children and Mothers in the Context of Trauma: An Initial Study of the Validity of the Child Reflective Functioning Scale'. *British Journal of Developmental Psychology*, 33(2): 203–217; Bizzi, F., Ensink, K., Borelli, J. L., Mora, S. C., and Cavanna, D. (2019). 'Attachment and Reflective Functioning in Children with Somatic Symptom Disorders and Disruptive Behavior Disorders'. *European Child & Adolescent Psychiatry*, 28(5): 705–717. A parallel distinction has been made in scaling the Parent Development Interview for mentalization by Suchman, N. E., DeCoste, C., Leigh, D., and Borelli, J. (2010). 'Reflective Functioning in Mothers with Drug Use Disorders: Implications for Dyadic Interactions with Infants and Toddlers'. *Attachment & Human Development*, 12: 567–585.

⁵ Thinking on embodied mentalizing has offered one way that questions about proto-thoughts and feelings are starting to be explored, though this has not seen much attention since Shai, D. and Fonagy, P. (2014). 'Beyond Words: Parental Embodied Mentalizing and the Parent Infant Dance', in M. Mikulincer and P. R. Shaver (eds), *Mechanisms of Social Connections: From Brain to Group*, Washington, DC: American Psychological Association, pp. 185–203. Another resource here may be Winnicott's concept of 'aliveness'. See Goldman, D. (2013). 'Vital Sparks and the Form of Things Unknown'. *Psychoanalytic Inquiry*, 33(1): 3–20.

⁶ E.g. Freud, A. (1977). 'Fears, Anxieties, and Phobic Phenomena'. *The Psychoanalytic Study of the Child*, 32(1): 85–90; Sandler, A. M. (1977). 'Beyond Eight-Month Anxiety'. *The International Journal of Psychoanalysis*, 58: 195–207; Sandler, A. and Freud, A. (1982). 'Discussions in the Hampstead Index on "The Ego and the Mechanisms of Defence": VI. Orientation of the Processes of Defence According to the Source of Anxiety'. *Bulletin of the Anna Freud Centre*, 5(1): 5–35. See also 'The problem of anxiety is a nodal point at which the most various and important questions converge, a riddle whose solution would be bound to throw a flood of light on our whole mental existence.' Freud, S. ([1916–1917] 2001). *Introductory Lectures on Psycho-Analysis* (standard edition), London: Hogarth Press, 15–16, p. 393.

therapy rather than a mentalization-based approach for children with anxiety.⁷ One reason for the relative lack of concern with anxiety for Fonagy and colleagues and other scholars working on mentalizing is that it does not align with their assumption that mental states are intentional (see Chapter 4). This assumption works well enough for some feelings like fear and disgust. But anxiety—and some other affects—can be free-floating, corresponding to a general sense of threat rather than a specific object. In general terms, anxiety can be conceptualized as the perception of present or future threat (general or specific), plus uncertainty (especially when extensive or poorly modulated), plus difficulty modulating uncertainty, in the absence of perceived help. The theory offered by Fonagy and colleagues is relevant to each of these elements. To sketch some illustrative directions for thinking about mentalization and anxiety: as we saw in Chapter 4, mentalization helps modulate threat appraisal and may help sustain access to help from others, whereas non-mentalizing reduces the modulation of uncertainty. In turn, high levels of anxiety hinders mentalizing. By contrast, reflection of the opacity of one's own or others' minds, or reflection on one's own or others' destructive feelings, may contribute to both anxiety and successful mentalizing. Drawing on the reflections in Chapter 5, it may be supposed that the modes of non-mentalizing contribute to relatively distinct and identifiable forms of anxiety:

- Pretend mode anxiety may be persecutory, because it is not tied down by actual states of affairs and the ordinary limitations of their ramifications.⁸
- Psychic equivalence anxiety may centre around the sense of crushing reality without alternatives or openings.
- Teleological mode anxiety may be seen as a sense of panicked urgency, in the desire to concretely resolve perceived problems and meet standards.

Further theory and research on the relationship between anxiety and mentalization may help contribute to further advancement of the paradigm, and to its clinical efficacy in planning and delivering treatment.

Some remaining questions: non-mentalizing

The most important issue facing the account of non-mentalizing, and perhaps the most significant limitation of the paradigm as a whole, is the lack of operationalization of pretend mode, psychic equivalence, and teleological mode. In our view, this has profoundly hindered the validation and advancement of these ideas over time. The attempt by another research group to operationalize the modes of non-mentalizing appears not to have been registered by Fonagy and colleagues.⁹ Whereas the modes of non-mentalizing have been central to the theorizing and advice on clinical technique offered by Fonagy, in fact this has all occurred without specific supporting evidence—and places the theory and clinical guidance subject

⁷ Fonagy, P. (2016). 'Guest Post: It's Important to Recognise When Childhood Anxiety is Becoming a Problem.' Accessed at: https://www.mumsnet.com/Talk/guest_posts/2710877-Guest-post-Its-important-to-recognise-when-childhood-anxiety-is-becoming-a-problem.

⁸ See e.g. Craib, I. (1994). *The Importance of Disappointment*, Bristol: Psychology Press.

⁹ Hausberg, M. C., Schulz, H., Piegler, T., Happach, C. G., Klöpper, M., Brütt, A. L., ... and Andreas, S. (2012). 'Is a Self-Rated Instrument Appropriate to Assess Mentalization in Patients with Mental Disorders? Development and first validation of the Mentalization Questionnaire (MZQ)'. *Psychotherapy Research*, 22(6): 699–709. See also Gagliardini, G. and Colli, A. (2019). 'Assessing Mentalization: Development and Preliminary Validation of the Modes of Mentalization Scale'. *Psychoanalytic Psychology*, 36(3): 249–258.

to the criticisms Fonagy levelled at psychoanalytic theory in the 1990s. If we compare the account of non-mentalizing to the six changes to psychoanalytic practice recommended by Fonagy as the basis for scientific credibility (see Chapter 2), this raises the following concerns:

- We do not actually know empirically that the three modes of non-mentalizing are negatively associated with mentalizing, or how strongly.
- We do not know that the reduction in non-mentalizing is associated with reduced symptoms or greater quality of life.
- We do not know whether MBT reduces all three modes of non-mentalizing. We do not know what components of MBT are effective for what forms of non-mentalizing.

A further set of questions for future work on non-mentalizing will be to consolidate an evolutionary perspective, attentive to mechanisms at individual and population levels. One way to do this would be to ask Tinbergen's four questions of each of the modes of non-mentalizing:¹⁰ How did it develop in the individual? What causes it? What is its function? How did it evolve in the species? Fonagy and colleagues have certainly addressed function and development. The account of the evolutionary basis of the modes of non-mentalizing is currently a sketch and warrants further attention. Cross-species comparison may help specify the role of representational and non-representational processes more sharply, alongside the phylogenetic building blocks of the non-mentalizing modes. However, perhaps most importantly, the proximal cause of non-mentalizing would benefit from exploration, hypothesis generation and empirical scrutiny. For instance, subsequent inquiry may consider forms of pleasure/enjoyment attending the modes of non-mentalizing, and help them to become reinforcing. It may be that non-mentalizing can feel invigorating, vital, or a relief in certain localized ways that help prompt this state of mind.¹¹ An obstacle that has repeatedly obscured effective examination of the cause of non-mentalizing has been the concept of 'trauma' as used by Fonagy and colleagues. Both recognition of the need for better specification and some of the ongoing problems with the concept of trauma in the work of Fonagy and colleagues can be seen in their recent chapter in the second edition of the *Handbook of Mentalising in Mental Health Practice*.¹² These include: 1) an overencompassing (and unoperationalizable) definition of trauma as 'experiences that go beyond the average expectable environment';¹³ 2) a characterization of the most important traumas as occurring, and healed, primarily by attachment relationships, without a specification of what an attachment relationship is; 3) the claim that 'adversity becomes traumatic only when it is compounded by the individual's sense that his or her mind is alone',¹⁴ a proposal that certainly contains insight, but has aspects of

¹⁰ Tinbergen, N. (1963). 'On Aims and Methods of Ethology'. *Zeitschrift Tierpsychology*, 20: 410–433.

¹¹ Cf. Bollas, C. (1996). 'Borderline Desire'. *International Forum of Psychoanalysis*, 5(1): 5–10. 'The borderline personality unconsciously seeks emotional turbulence because this complex of affect is the shape of the object of desire. Whether these people were intrinsically disturbed as infants, or, whether the early object world was itself disturbing, they knew the maternal object as disruptive effect. This effect then became the shape of the object, so, in seeking turbulence they are in fact constituting the primary object. As painful and disturbing an event as this is, it is nonetheless desired and finding themselves in states of distress is unconsciously gratifying. This person cultivates "borderline objects" which evoke turbulent frames of mind' (p. 6).

¹² Luyten, P., Malcorps, S., Fonagy, P., and Ensink, K. (2019). 'Mentalising and Trauma', in Anthony Bateman and Peter Fonagy (eds), *Handbook of Mentalising in Mental Health Practice* (2nd edn), Washington, DC: American Psychiatric Association, pp. 79–102.

¹³ Ibid. 82.

¹⁴ Ibid. 86.

a soundbite. It seems to assume that reframing of one's experience in discussion with others is the sole means by which adversity does not become traumatic. This claim is difficult to evaluate given that the term 'trauma' is used so generally, and given that the account of 'with one's mind' seems to slide about between metaphor, phenomenological description, and causal mechanism.

Questions remain also regarding the status of 'alienation of the alien self'. Has the concept dropped out of the theory? Or is it in fact a fourth mode of non-mentalizing? At present, links between this concept and theoretical developments of the 2000s and 2010s (e.g. epistemic trust, the p-factor) remain poor. This is despite the strong potential for relevant connections—for instance, between the alien self and epistemic credulity. If it is decided that alienation of the alien self is to be rehabilitated as a live element of the theory, a question that could be valuable to pursue is whether there are forms of alien self that contribute positively to mentalizing under certain conditions. For instance, it might be considered that the ambitions for ourselves to which we aspire form a kind of 'alien self', which may contribute to rumination but which may also, or instead, contribute to efforts to achieve a life commensurate with our hopes and values.¹⁵ If the 'alien self' is reinvigorated as a concern, another question might be whether alien selves with the same correlates and clinical meanings are created through sustained use of pretend mode, psychic equivalence, and teleological mode. There are hints in the work of Fonagy and colleagues that the different forms of non-mentalizing populate different kinds of innerscape, but this has not been considered explicitly. For instance, Fónagy and Fonagy's account of primary and secondary meanings (see Chapter 2) may suggest that teleological mode focuses on primary meanings at the expense of secondary meanings, and the intentional at the expense of the non-intentional, and so loses the richness and nuance these provide.

Finally, an exciting set of questions is opened up by the emerging ideas of Fonagy and colleagues regarding collective non-mentalizing. Among other things, their account of systematically distorted public communication undermining the deliberative basis of democratic societies can be regarded as an incipient political theory.¹⁶ This would be attended by an incipient theory of the psychosocial mechanisms of dehumanization. For instance, the political objectification of humans may be analysed in terms of how propaganda is sustained by pretend mode, how the emotive construction of threats is sustained by psychic equivalence, and how the replacement of the intentions of individuals with reifications of their labour or political 'will' is sustained by teleological mode.

Finally, as we saw in Chapter 9, Fonagy and colleagues identify a major shift in their current thinking on the basis of attention to the wider social system. We perceive great potential benefits of a 'distributed model' of mentalizing—and epistemic trust—rather than treating these as capacities lodged solely in the heads of individuals. This would help facilitate attention to the conditions that support mentalizing capacity and forms of epistemic trust. For instance, attention to the material conditions of mentalizing could open opportunities for dialogue and collaborative research with researchers in architecture, urban planning, and geography.

¹⁵ Cf. Blos, P. (1974). 'The Genealogy of the Ego Ideal'. *The Psychoanalytic Study of the Child*, 29(1): 43–88.

¹⁶ See Habermas, J. (1970). 'On Systematically Distorted Communication'. *Inquiry*, 13(1–4): 205–218.

Some remaining questions: epistemic trust

Mentalizing and non-mentalizing has been a long-standing concern for Fonagy and colleagues, and the outstanding questions above arise in reflection on an established research programme and clinical modality. By contrast, their work on epistemic trust is more recent and emergent. We hope that future work will include efforts to operationalize epistemic trust, epistemic mistrust, epistemic vigilance, epistemic hypervigilance, and epistemic credulity. However, there are some theoretical questions that may have precedence. This includes the question of the object of 'trust'. Their account tends to imply that trust is in the thoughts of others. It remains ambiguous whether they think it is possible to have epistemic trust or mistrust in the feelings of others. The capacity for epistemic trust or mistrust in the thoughts or feelings of oneself would also benefit from further attention, because this has barely been acknowledged but would seem to be of the utmost importance for linking the idea of epistemic trust to the capacity to reconceive thoughts and feelings as part of mentalizing.

It would be good to see attention to the relationship between epistemic mistrust and the forms of non-mentalizing. An obstacle to date has been the rather limited diet of examples of epistemic trust and mistrust used by Fonagy and colleagues, which focus on parent–infant and therapist–client interactions. Both these forms of interaction have quite specific power dynamics that are unrepresentative of most adult interactions, and which contribute to unusual configurations of trust and mistrust. One consequence, discussed in Chapter 9, is that the role of power dynamics in epistemic trust and mistrust can be left out of frame by Fonagy and colleagues. This has left significant open questions about how the epistemic mistrust of others, including in collective forms through epistemic oppression, can contribute to forms of non-mentalizing. This could help advance the interest Fonagy and colleagues have demonstrated in identifying the societal conditions that contribute to the prevalence of borderline personality disorder. Their current emphasis on the role of 'fragmented, dehumanizing social systems' across society runs some risk of aligning with Fonagy's pessimism about social change. By contrast, attention to the ways that epistemic mistrust contributes to institutionalized forms of non-mentalizing may offer a more specific, testable, and optimistic locus for concern.¹⁷

Core strengths of the theory

Having described some particular reflections across the chapters of the book which we hope will be constructive for the paradigm, and identified some outstanding questions for future theory and research, we would like to close by highlighting what we see as core propulsive strengths of the theory of mentalization and epistemic trust. In closing the present book, we wish to draw attention to five features of the paradigm that we see as contributing especially to its tremendous strengths.

First, as documented in Chapter 2, Fonagy and colleagues have assembled a remarkable institutional and theoretical edifice, permitting the movement of knowledge across different domains. Perhaps most remarkably, their paradigm has helped sustain bi-directional transfer

¹⁷ This would align with advances in sociological theory. The concept of general societal fragmentation, which Fonagy and colleagues take from Durkheim, is generally regarded by sociologists today as limited; more specific causes tend to be identified in the rules, norms and routines that organise institutionalised practices. Baert, P. and Da Silva, F. C. (2010). *Social Theory in the Twentieth Century and Beyond*, Cambridge: Polity, Chapter 1.

of knowledge between psychoanalysis and health services research and policy. Fonagy and colleagues have been unusually reflexive in their work, repeatedly sensing earlier than peers which way the winds are blowing in sociological pressures faced by psychological research and practice. As well as helping defend against threats and make use of opportunities represented by these pressures, this reflexivity has also helped their work in other ways—for instance, in helping them tailor their message to different constituencies.

Second, as Chapter 3 identified, their theory is intrinsically and insistently developmental.¹⁸ This has contributed to the acuity of research, such as the Anna Freud longitudinal study (see Chapter 2) and the pre-natal use of the Adult Attachment Interview (see Chapter 3). It has helped the theory incorporate insights from the developmental psychopathology movement, such as the p-factor and the idea of symptom networks (see Chapter 7). The developmental perspective has also underpinned adaptations of MBT to support children and adolescents, and parents and other caregivers (Chapter 8). Without a doubt, the theory of mentalization has contributed to a great deal of important empirical work with relevance across a range of disciplines.

Third, MBT offers a clinical modality with impressive transferability between settings and evidence of efficacy for many populations. Early indications suggest that the modality has adapted very fluently to remote delivery during the COVID-19 pandemic. Some elements of MBT are relatively generic, and the training burden for practitioners is comparatively light. However, MBT also offers compelling and distinctive clinical guidance—for instance, in the proposal that therapists pursue ‘contrary moves’ to help expand patients’ repertoire of mentalizing (Chapter 8). Likewise, there is a great deal of clinical acuity in the account of the modes of non-mentalizing. The distinction between hypermentalizing and mentalizing is important in its commitment to the therapeutic value of accurate integration objectively perceived with subjectively conceived reality. For all that they have not been operationalized and so remain scientifically underdeveloped, the forms of non-mentalizing are perspicacious characterizations of ways in which information may be filtered, obstructing generative doubt. They represent a distinctive contribution, irreducible to either existing cognitive models or psychoanalytic concepts of defence.

Fourth, the ideas of Fonagy and colleagues have an underpinning architecture, reflected in the 18 key concepts explored in this book. The writings of Fonagy and colleagues are carefully crafted to be accessible to specific audiences, and their clarity can give the impression that they need only be read once to be fully understood. This has misled some critics, particularly from within psychoanalysis, who misapprehend Fonagy and colleagues as a simplified and superficial version of familiar psychoanalytic notions. Yet, when the works of Fonagy and colleagues are looked at in the round, an underpinning architecture comes into focus, and reveals the huge questions pursued by the researchers and, ultimately, the depth of their response. So for instance a measure such as the Reflective Functioning Questionnaire may appear to be a simple set of items to a psychoanalytic critic, but in fact possesses ingenious and theoretically-innovative aspects in the conceptualization of certainty, uncertainty, and mental health. In one of his papers, Derrida asks, ‘What is it in a “great” work, let’s say of Plato, Shakespeare, Hugo, Mallarmé, James, Joyce, Kafka, Heidegger, Benjamin, Blanchot, Celan, that resists erosion? What is it that, far from being exhausted in amnesia, increases its reserve to the very extent to which one draws from it, as if expenditure augmented the capital?’¹⁹ In the case of Fonagy and colleagues, a strength of their work is its

¹⁸ See Sroufe, L. A. (2009). ‘The Concept of Development in Developmental Psychopathology’. *Child Development Perspectives*, 3(3): 178–183.

¹⁹ Derrida, J. (1989). ‘Biodegradables’. *Critical Inquiry*, 15: 812–873.

underpinning architecture, which allows clinical, theoretical, and research activities to cross-fertilize and sustain one another. For instance, ideas from attachment theory about adaptation have been drawn upon to qualify the potential pathologization of patients associated with diagnosis and deficit models; and empirical findings from developmental psychopathology (e.g. symptom networks) have been used to qualify limitations with attachment theory.

Fifth, something frankly quite distinctive about the work of Fonagy and colleagues has been their openness to development and refinement of their theory, through the incorporation of ideas from different communities. Theory, research, and clinical practice have been conducted in a way that, whenever possible, maintains ‘inquisitive stance’: acknowledging the potential for influence, surprise, change, and learning with and from others. Key moments include:

- The integration of ideas from within psychoanalysis: the Anna Freudian and Kleinian traditions of psychoanalysis, as well as the work of Winnicott (Chapters 1, 2 and 6).
- The integration of ideas from within cognitive science: work by researchers on social cognition (Chapter 3) and Gergely and Csibra’s natural pedagogy theory (Chapter 7).
- The integration of ideas from within developmental science: the tradition of empirical attachment research in the 1990s (Chapter 3), and subsequent advances in developmental psychopathology and evolutionary approaches to behaviour (Chapter 7).
- A significant transition seems to be presently underway in growing concern with culture and the wider social system (Chapter 9).

It is our hope that the reflections offered in the present book may contribute to greater understanding of the work of Fonagy and colleagues. We also hope they will prove useful for further refinements of the paradigm.

Index

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.

- adaptation 167–77, 210–11
- Adult Attachment Interview 58, 59–61
- affective mentalizing 83–84
- aggression 161–65
- Ahmed, S. 273
- Ainsworth, M. D. S. 59, 170–72
- alcohol use 213–15
- alien self 151–57
 - aggression 161–65
 - externalization 132, 154–57, 289
 - sexuality 157–61
- Allison, E. 9, 91–92, 128, 142–44, 175–76, 179–80, 182–83, 230–31, 240, 241–42, 249–50, 272, 276
- Allport, G. W. 138
- alpha function 61, 105–6
- Althusser, L. 275–76
- Anna Freud Centre 15–17
 - Contact and Residence Disputes service 39–40
 - current services 39–42
 - Fonagy as research director 26–35
 - On My Mind resources 41–42
 - Parent–Child Project 39–40, 60–61
 - retrospective study 35–38
 - transformations 38–48
 - Young Champions 41–42
- anthropological culture 266–67
- antisocial personality disorder 164–65
 - mentalization-based therapies 236–37
- anxiety 36–37, 47, 48, 156–57, 204, 219–21
 - positive prognostic indicator 36–37, 286–87
- arousal 87–89, 117–18, 121, 195–96
- Asen, E. 193–94, 260, 262–68, 269, 272, 277–78
- attachment 51, 58–69
 - adaptation 167–77, 210–11
 - borderline personality disorder 81–82, 87–88
 - disorganized 59–60, 69–76, 90, 115, 135–36, 182–83, 185–86
 - epistemic trust 182–91
 - mentalizing 67, 78, 87–88
 - ostensive cues 181–82
 - reflective function 61–64
 - revision of theory by Fonagy and colleagues 64–69
 - strategies and transactions 211
- autism 51, 57
- automatic mentalizing 85–87
- Bakermans-Kranenburg, M. J. 189–91
- Baron-Cohen, S. 51, 54–55
- Bateman, A. 8, 21, 50–51, 68, 81–82, 99, 104–5, 107–8, 111, 123, 147, 149, 150, 189, 195–96, 197–98, 201–2, 217–18, 222–25, 236–37
- Baudelaire, C.-P. 114
- Belsky, J. 173–74, 211
- Berthelot, N. 90
- Bevington, D. 238–39
- Bion, W. R. 56–57, 105–6, 133–34
- Bo, S. 189
- Bollas, C. 44–45, 61
- borderline personality disorder (BPD) 21–22, 50–52, 91–92
 - adaptation 175–76
 - attachment 81–82, 87–88
 - defending the diagnosis 197–98
 - dehumanizing social systems 270–72
 - emergency form of social understanding 177
 - epidemiology 271–72
 - external and internal mentalization 82–83
 - hypermentalizing 109–10
 - indicator of general psychosocial difficulties 201–2
 - mentalization-based therapies 217–18, 231–32
 - mental states 53–54
 - psychic equivalence 117
 - therapeutic focus on mentalization 68–69
 - trauma 81–82
- borderline state 16–17
- Bouchard, M. A. 231–32
- Bowlby, J. 51, 196–97
- Brent, B. K. 205
- bullying 236
- Calkins, M. W. 136–37
- Campbell, C. 9, 91–92, 175–76, 193–94, 195–96, 201, 230–31, 241–42, 249–50, 262–68, 272, 276, 277–78
- Carlson, E. 75
- Caspi, A. 200–1
- Cassidy, J. 75
- Chiesa, M. 8
- child abuse 37–38, 52, 65, 90–91, 111–12
- Choi-Kain, L. W. 80
- cognitive behavioural therapy (CBT) 30, 57–58, 122–23
- cognitive informational culture 266–67
- cognitive mentalizing 83–84
- coherence 58–59, 62, 147, 148, 149–50, 153–55
- collective mentalizing capacity 256–62
- collective non-mentalizing 260, 269, 276–77, 289
- Colli, A. 122–23
- concepts 3–6, 49
- concrete meaning 12
- constitutional self 148
- contracts 122–23
- contrary move 226–27
- controlled mentalizing 85–87
- Corriveau, K. H. 182–83
- COVID-19 268–69
- creativity 113–14
- critical thinking 193

- Crittenden, P. 68, 172, 210–11
 Csibra, G. 120, 180–81
 culture 258, 262–81
 cycle of abuse 65
- Debbané, M. 205
 defences 29–30, 168–69
 De Meulemeester, C. 232–33
 depression 36–37, 46–47, 48, 204, 212–13, 219–21, 234–35, 286–87
 developmental help 16–17, 27–28
 diagnoses 196–208
 disorganized attachment 59–60, 69–76, 90, 115, 135–36, 182–83, 185–86
 dissociation 115
 Di Stefano, G. 256–57
 distorted social cognition 276
 Dotson, K. 274–75
 Dunn, J. 54–55
 Durkheim, E. 271
 dynamic interpersonal therapy 219–22
 dysregulated emotion 117–18
- eating disorders 149
 Education for Wellbeing trial 253–54
 Egeland, B. 65
 emotion regulation 110, 117–18
 empathy 17–18, 57, 83–84, 175, 237
 emptiness 24, 54
 Ensink, K. 90–91, 101, 111–12, 131
 epistemic mistrust 179, 183, 188–89, 195, 201, 211, 214, 290
 persistent 191–93
 epistemic trust 177–96
 absence of 191–93, 194–95
 attachment 182–91
 core strengths of theory 290–92
 mentalization 191, 195–96
 opportunities for future work 290
 epistemic vigilance 177–79, 182–83, 185–86, 192, 193–94, 195–96, 242–43, 273
 evidence-based health care 42–43, 46–47
 externalization of alien self 132, 154–57, 289
 external mentalizing 82–83, 99
- Fairbairn, W. R. D. 144
 family therapy 260
 Fearon, P. 78–79
 flashbacks 117
 Fónagy, I. 11–13, 15–16
 Fonagy, P.
 attachment theory revisions 64–69
 #BlackLivesMatter campaign 248–49
 central collaborators 7–9
 on concepts 4, 49
 constitutional self 148
 criticism of treating self as an entity 138–41
 evidence-based practice 42–43, 46–47
 family background 11–15
 first use of ‘mentalization’ 56
 humour use 21
 importance of friendships 242, 243–44
 manual of techniques for child psychoanalysis 27–28
 mentalization versus theory of mind 56–57
 pessimism about social change 247–48
 policy engagement 45–46, 48, 254–56
 in psychoanalysis with Hurry 15–21, 57
 as research director at Anna Freud Centre 26–35
 scientific evaluation of psychoanalytic practice 31–35, 43, 46–47
 self-criticism 77–82
 summary of career 1
 on targeting resources 247
 at University College London 26
 on violence 163
 What Works for Whom? 42–43
 as a young analyst 21–25
- Fonagy’s clinical cases
 Billy 253
 Callum 156–57
 David 162
 Jessie 18
 Mr S 21–25, 52
 Mr T 24–25
 Rebecca 105
 William 27
- Foucault, M. 275–76
 fragmentation 142–45, 159, 163
 Freud, A. 16, 28–30, 168–69, 197
 friendships 241–45
 Frith, U. 85–86
 Fuggle, P. 238–39
 full mentalizing 98–99
- Gagliardini, G. 122–23
 gender dysphoria 198–99
 gentle touch 180–81
 Gergely, G. 8, 66, 81, 120, 180–81
 Gibbon, L. 206–8
 Gillespie, S. M. 237
 Goldberg, S. 200
 Green, A. 142
 Grossman, W. I. 145
 group-based MBT 225, 240, 260–62
 Group for the Study of Borderline
 Cases 16–17, 24
 Gunderson, J. G. 80
- Harris, P. I. 182–83
 Hartmann, H. 136–37
 Havsteen-Franklin, D. 86–87
 Hegel, G. W. F. 139–40
 Hesse, E. 59–60
 Higgitt, A. C. 7, 42–43, 56–57, 60–61, 71, 138–39, 209–10, 246
 Hill Collins, P. 274–75
 Holmes, J. 167–68
 homelessness 240
 humour 12, 20–21
 Hurry, A. 15–21, 57
 Hutchins, E. 2–3
 hyperactive reflective function 63
 hyperarousal 107–8, 174–75
 hypermentalizing 63, 103–4, 106–10, 133, 205, 235–36, 269–70, 291
- imaginative capacity 111, 112–13
 incoherence 142–45, 146, 159, 163

- inquisitive stance 21
- insight (cognitive) 17–18, 19
- integrity of self disorders 149
- intensity of therapy 36–37, 38
- intention/intentionality 97–98, 143–44, 280–81
- internal mentalizing 82–83, 99
- internal observer 61
- internal working model 66
- interpersonal interpretive function 78
- interpersonal transaction cycles 212–14
- intersubjectivity 113–14, 180–81

- Jacobson, E. 141
- Jacobvitz, D. 65
- Jones, P. B. 204–6

- Karterud, S. 240, 260–61
- Kernberg, O. F. 167–68, 201–2
- kindness 180–81
- Kongerslev, M. T. 189, 237

- language
 - Adult Attachment Interview 60–61
 - concepts and 6
 - Ivan Fónagy's interest in 11–13
 - primary and secondary meanings 13–14
 - psychological research and theory 6
- Lecours, S. 231–32
- Lemma, A. 106, 193–94, 219–20
- Luyten, P. 8–9, 80–81, 82–92, 106, 149, 175–76, 193–94, 212–14, 241–42, 244
- Lyon, F. 191

- Main, M. 58–60, 69–70, 75, 171–72
- marked mirroring 66
- Marty, P. 56
- meaninglessness 142–45, 159, 163
- Meins, E. 182–83
- membership cases 21–25
- mentalization-based therapies 216–37, 291
 - advocacy for patients 245–46
 - antisocial personality disorder 236–37
 - balance 226–27
 - borderline personality disorder 217–18, 231–32
 - contrary move 226–27
 - development 216–19
 - dynamic interpersonal therapy 219–22
 - effectiveness 231–33
 - family therapy 260
 - group-based 225, 240, 260–62
 - growing popularity of 222–25
 - individual 225, 239
 - latest manual 225
 - mentalizing functional analysis 228–29
 - ostensive cues 229–31
 - potential for harm 233–37
 - psychoeducation 226
 - self-harm 233
 - stop and re-wind 229
 - symptom specification 226
 - trauma 227–28
- mentalizing/mentalization
 - affective 83–84
 - affect regulation 77–78
 - attachment and 67, 78, 87–88
 - attentional control 77–78
 - automatic 85–87
 - changes in meaning 49–50
 - characteristics of good mentalizing 99–101
 - as a clinical technique 57–58
 - cognitive 83–84
 - cognitive capacities 81
 - collective mentalizing capacity 256–62
 - confabulation 152–53
 - controlled 85–87
 - core strengths of theory 290–92
 - culture 258, 262–81
 - definitions 92–101
 - disorganization 69–76
 - epistemic trust 191, 195–96
 - external 82–83, 99
 - four polarities of 82–92, 125–26, 286
 - full mentalizing 98–99
 - genetic factors 78–79
 - habit 274–75
 - hierarchical organization 80–81
 - internal 82–83, 99
 - as a means to an end 196
 - mentalization mafia 6
 - mentalizing functional analysis 228–29
 - mentalizing profile 89–90
 - mentalizing quadrant 86–87
 - multidimensional assessment 89–90
 - not mentalizing and 102
 - opportunities for future work 285–87
 - origins of the concept 50–58
 - others 84–85, 101
 - reflective function and 62–63, 80
 - resilience 128–29
 - self 84–85, 101
 - self-report scale 129–30
 - social experience 79
 - specific mentalizing 98–99
 - synthesis of experience and understanding 127–31
 - theory of mind and 56–57
 - therapeutic focus in borderline personality disorder 68–69
- mental states 52–54, 55, 133–35, 180–81
- metaphor
 - meaning of language 12
 - in science 31–32
- mood 134–35
- Moran, G. S. 7, 60–61, 161–62, 216–17, 242

- neoliberalism 43–45
- Nerval, G. 114
- Neufeld, S. 204–6
- New Public Management 269–70
- Nolte, T. 206–8
- non-mentalizing 102–32, 167–68
 - collective 260, 269, 276–77, 289
 - evolutionary perspective 288–89
 - externalization of alien self 132, 154–57, 289
 - hypermentalizing 63, 103–4, 106–10, 133, 205, 235–36, 269–70, 291
 - not mentalizing and 102
 - opportunities for future work 287–89
 - pretend mode 103–16, 269–70

- non-mentalizing (*cont.*)
 psychic equivalence 103, 116–19
 systemic obstacles to mentalizing thoughts or feelings 132
 teleological mode 103, 119–23, 269–70
 withdrawal 131–32
- Normandin, L. 111–12
- ostensive cues 180–82, 229–31
- Patalay, P. 202
- Payne, C. 48
- Pears Family School 259–60
- perspective-taking 112, 114–15
- p-factor 200–8, 209–12, 213–14
- playful exploration 111, 113
- pleasure principle 52
- Polek, E. 206
- policy issues 45–46, 48, 254–56
- political solicitation 187–88
- post-traumatic stress disorder 172, 174–75
- pretence 12
- pretend mode 103–16, 269–70
- prevention 246–56
- primary unconscious 142–50
- projective identification 154–55, 180–81
- psychic equivalence 103, 116–19
- psychoanalytic unconscious 143–45
- psychoeducation 226
- public mental health 246–56
- queer theory 272–76
- Reeves, J. K. 38–39
- reflective function 58–69, 80, 101, 181–82
- Reflective Functioning Questionnaire (RFQ) 129–31, 207–8
- reflective self 61
- representational violence 163
- resilience 29–30, 128–29, 268–69
- responsibility 278–81
- Rizq, R. 235
- Rogers, C. 136–37
- romantic love 118–19
- Rossouw, T. I. 233
- St Clair, M. C. 204–6, 213–14
- Sandler, J. 3–4, 141, 282
- school settings 249–56, 258–60
- scientific evaluation, psychoanalysis 31–35, 43, 46–47
- secondary gain 4, 282
- security 68
- self 133–65
 alien *see* alien self
 coherence 149–50
 concept 136–41
 constitutional 148
 learning about 184–85
 mental disorders 149
 mentalizing 84–85, 101
 primary unconscious 142–50
 reflective 61
 self-structure destabilisation 150
 self-concept disorders 149
 self-denigration 187–88
 self-harm 47, 149, 162–63, 233
 self-image disorders 149
 self-management 43–45, 246
 self-recognition disorders 149
 self-regulation 43–44, 246, 247–48, 254
 sensitive care 64–65, 112, 183–84
 sexual abuse 90–91, 111–12
 sexuality 118–19, 157–61
 Sharp, C. 107–8, 109–10, 201–2, 237
 Skårderud, F. 115–16, 122–23, 149, 211–12
 social cognition 54–55, 63–64, 276
 social denigration 187–88
 social environment 177, 239–46
 social understanding 23–25
 Solomon, J. 69–70
 specific mentalizing 98–99
 Sperber, D. 177–79
 Sroufe, A. 75
 Sroufe, L. 65
 Steele, H. 7, 58, 60–61, 62
 Steele, M. 7, 58, 60–61, 62
 stop and re-wind 229
 Strange Situation 59, 60–61, 78–79, 182–83
 substance use 213–15
 Suchman, N. E. 234–35
 symptom contexts and networks 209–15
- talking therapies 119
- Target, M. 7–8, 33, 35–38, 46–47, 57–58, 67, 69–70, 78–79, 80–82, 104, 106, 112, 113, 114–16, 119, 120, 124–25, 140, 148, 152–53, 155, 156–63, 193–94, 219–20, 231–32, 235, 258
- teleological mode 103, 119–23, 269–70
- terminology 3–6
- Tessier, V. P. 111–12
- theory of mind 51, 56–57
- THRIVE model 252–53
- training, child psychotherapy 40–41
- trauma 53, 81–82, 117
 mentalization-based therapies 227–28
- Twemlow, S. W. 244, 250–52
- twin studies 78–79
- unconscious
 primary 142–50
 psychoanalytic 143–45
- Van Harmelen, A. I. 242–43
- Van Ijzendoorn, M. H. 189–91
- Vincent, S. 278
- violence 163
- Winnicott, D. W. 111, 127
- withdrawal 131–32
- Wolpert, M. 252
- working memory 228